Title
Essays on Cross-sectional Methods in Macroeconomics

ISBN
9798186182658

Author
Donaldson, Maria Paula

Publication Date
2026-08-04

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Cross-sectional Methods in Macroeconomics

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Paula Donaldson

Committee in charge:

Professor Johannes Wieland, Chair
Professor Juan Herreño
Professor Nir Jaimovich
Professor Munseob Lee
Professor Valerie Ramey
Professor Fabian Trottner
Professor Steve Wu

2026

The Dissertation of Paula Donaldson is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2026

DEDICATION

To my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor, Johannes Wieland, for his guidance and support throughout my PhD. His mentorship and generosity shaped me as a researcher, and his example inspires the rigor, curiosity, and humility to which I aspire. I am indebted to him for everything he has taught me.

I am also thankful to my committee members, Juan Herreño, Nir Jaimovich, Munseob Lee, Valerie Ramey, Fabian Trottner and Steve Wu, for their time, support, and invaluable feedback. I would like to extend a special thank you to Juan and Fabian. Their kindness and support made my PhD a much more enjoyable experience, and I am deeply grateful to them.

My deepest gratitude goes to my family for always being there for me. To my parents, thank you for your unconditional love and for always supporting my decisions, even when they took me far from home. To my sisters, thank you for your love and your *complicidad*. I could not have done this without you.

Moving to a new country and pursuing a PhD would have been a very different experience without the friendships I made along the way. I would especially like to thank Fulvia, Jordan, Francesco, and Andre. We went through this process together, sharing ups and downs, endless conversations, and many beautiful moments over the years. I will always be grateful for the friendships that accompanied me through these years.

I would also like to thank my coauthors. Working with them taught me a lot and made research both fun and fruitful.

Finally, I would like to thank my former colleagues and mentors at the Central Bank of Argentina for sparking my interest in research. I am especially grateful to Mauro Alessandro, whose encouragement and confidence in me gave me the push I needed to pursue a PhD.

Chapter 1, in full, is based on the paper "Cross-sectional Identification with Heterogeneous Exposure to General-equilibrium Effects" currently being prepared for submission. The dissertation author is the sole author of this paper.

Chapter 2, in full, is based on the paper from Federico Bernini, Paula Donaldson, Ezequiel

Garcia-Lembergman, and Leticia Juarez, "The Financial Channel of Tax Amnesty Policies" currently being prepared for submission. The dissertation author is one of the primary authors of this paper.

Chapter 3, in full, is based on the paper from Paula Donaldson Paula and Johannes Wieland, "Why are Some Recoveries Weak and Others Strong?" currently being prepared for submission. The dissertation author is one of the primary authors of this paper.

VITA

2014	Bachelor of Arts in Economics, Universidad de Buenos Aires.
2015-2020	Senior Researcher, Economic Research Division, Central Bank of Argentina.
2021-2026	Teaching Assistant, University of California San Diego.
2021-2026	Graduate Student Researcher, University of California San Diego.
2026	Doctor of Philosophy, University of California San Diego.

FIELDS OF STUDY

Primary Field: Macroeconomics. Secondary field: International Economics.

ABSTRACT OF THE DISSERTATION

Essays on Cross-sectional Methods in Macroeconomics

by

Paula Donaldson

Doctor of Philosophy in Economics

University of California San Diego, 2026

Professor Johannes Wieland, Chair

This dissertation examines how cross-sectional methods can be used to study the transmission of macroeconomic shocks, combining methodological contributions with applications to fiscal policy, financial intermediation, and business cycle dynamics.

In Chapter 1, I study cross-sectional designs that exploit heterogeneous exposure to aggregate shocks for estimating partial-equilibrium elasticities in macroeconomics. I show that these designs can fail to identify partial-equilibrium elasticities when exposure to the shock of interest is correlated with exposure to aggregate variables that move in general-equilibrium. I develop a test for this identification failure and propose a new decomposition method that leverages both cross-sectional and time-series variation to address it. Applying the method to

estimate US cross-sectional fiscal multipliers, I find that accounting for heterogeneous exposure to monetary policy reduces the estimated two-year multiplier from 1.5 to 1.

In Chapter 2, we study how asset repatriations can deepen financial systems, using Argentina's 2016 Tax Amnesty as a case study. The program generated a large inflow of foreign-currency deposits into banks, allowing us to study the transmission of a currency-specific liquidity shock. Exploiting variation in banks' exposure to this inflow and firms' exposure to banks, we find positive effects on firms' financial and real outcomes. We discipline a small open economy model with heterogeneous banks and firms using our reduced-form elasticities and use it to quantify aggregate effects and the roles of currency composition and openness in transmission.

In Chapter 3, we use regional heterogeneity across U.S. states to study why some national recessions are followed by stronger recoveries than others. Leveraging heterogeneous exposure to the national business cycle across U.S. States, we estimate the trajectory of more exposed U.S. States relative to less exposed ones. For the 1990-1, 2001, and 2007-9 recessions we estimate that more exposed States experienced a stronger boom-bust cycle and for the other recessions we estimate a deeper V-shaped recession in more exposed States. Our estimates support theories that these recessions are caused by different shocks. In a quantitative model matched to our cross-sectional estimates, boom-bust cycles persistently depress the natural rate of interest r^* in the recovery.

Chapter 1

Cross-sectional Identification with Heterogeneous Exposure to General-equilibrium Effects

1.1 Introduction

Cross-sectional methods are increasingly used in macroeconomics, international and spatial economics to study how the economy responds to aggregate shocks. When households, firms, or regions differ in their exposure to these shocks, a comparison of their relative responses recovers the elasticity of the cross-sectional outcome to the aggregate shock. Examples include fiscal, trade, and monetary policy, among other economic settings in which heterogeneous exposure to aggregate shocks can be exploited for identification.[1]

These methods are designed to identify partial-equilibrium, or micro-level, elasticities. This requires separating the direct effect of the shock from the aggregate effects it sets off in general equilibrium. The typical empirical strategy combines time fixed effects, which absorb the aggregate effects, with cross-sectional differences in exposure, which provide the variation needed to isolate the direct effect. Throughout this paper, I refer to these cross-sectional elasticities as *portable*[2] elasticities, echoing the terminology introduced by Nakamura and Steinsson (2018).

[1] Some influential examples include fiscal and tax policy (Parker et al., 2013; Nakamura and Steinsson, 2014; Zidar, 2019; Pennings, 2021), credit supply (Mian and Sufi, 2014), monetary policy (Ottonello and Winberry, 2020; Peydró et al., 2021), trade shocks (Autor et al., 2013, 2021), investment elasticities (Zwick and Mahon, 2017), and wealth effects (Guren et al., 2021b; Chodorow-Reich et al., 2021), among others.

[2] I use the term *portable* to describe elasticities estimated at different levels of aggregation that capture the

Such elasticities are valuable for their external validity and their role in policy evaluation and counterfactual analysis. By isolating well-identified responses to shocks, they enable progress on important questions that aggregate time series analysis alone cannot resolve.

But this strategy rests on a strong identifying assumption: that exposure to the shock is uncorrelated with sensitivity to aggregate variables that co-move in general equilibrium. In practice, the same characteristics that drive exposure to one policy often also shape sensitivity to others. For example, regions more exposed to fiscal expansions may also be more or less sensitive to interest rates, which typically respond to fiscal policy. In such settings, cross-sectional variation in exposure does not isolate the direct effect of the shock. Identification fails because the estimated elasticity is contaminated by heterogeneous responses to dynamic general equilibrium effects. I refer to such environments as HEDGE settings, short for *heterogeneous exposure to dynamic general equilibrium effects*. This paper asks: How can portable elasticities be reliably identified in such environments?

I provide a new framework that makes it possible to identify portable elasticities in HEDGE economies. The approach combines elements from cross-sectional and time-series analysis to decompose the elasticity estimated using the typical two-way fixed effects (TWFE) design into the portable elasticity of interest and a HEDGE term. I use the framework to estimate cross-sectional fiscal multipliers in the US, allowing states to differ in their exposure to changes in both defense spending and interest rates. The two-year multiplier drops from 1.5 to 1 when applying the decomposition. This finding challenges the view that TWFE estimates of cross-sectional fiscal multipliers are independent of the monetary stance. The estimated portable multiplier is smaller for two reasons: (i) US states that are more exposed to defense spending are also less sensitive to monetary policy shocks, and (ii) monetary policy responds to an expansionary defense spending shock by tightening interest rates. Finally, I develop a two-region TANK model, combining features of Nakamura and Steinsson (2014) and Herreño

cross-sectional, direct effect of an aggregate shock while holding other aggregates fixed. Depending on the context, this may correspond to an individual optimization problem (e.g., marginal propensities to consume) or to a partial-equilibrium elasticity where some local market clears (e.g., cross-sectional multipliers).

and Pedemonte (2025), to interpret the empirical findings. The model analogue of the TWFE multiplier can vary sharply across monetary regimes once HEDGE is present, challenging its usefulness for bounding the aggregate multiplier. In contrast, the portable multiplier maintains its ability to provide upper and lower bounds on the aggregate effects of fiscal policy shocks across monetary regimes.

The Identification Challenge

First, I establish the conditions under which HEDGE represents an identification challenge to elasticities estimated using the following TWFE specification:

$$Y_{it+h} = \beta_h s_{ix} \varepsilon_{xt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{1.1}$$

where Y_{it} is the outcome of unit i at time t (e.g., local GDP), s_{ix} denotes the exposure of unit i to an aggregate variable X, and ε_{xt} is an innovation to X in period t. Consider an economy[3] where units differ in their exposure to two aggregate variables: G_t, the variable of interest (e.g., government spending), and R_t (e.g., monetary policy rate). Then it can be shown that the estimate of β_h from a TWFE regression like (1.1) converges to:

$$\hat{\beta}_h \rightarrow \beta_h^P + \gamma \underbrace{\frac{cov\left[s_{ig}, s_{ir}\right]}{v\left[s_{ig}\right]}}_{\phi} \underbrace{\frac{cov\left[R_{t+h}, \varepsilon_{gt}\right]}{v\left[\varepsilon_{gt}\right]}}_{\boldsymbol{R}_{g,h}}, \tag{1.2}$$

where ε_{gt} is an innovation to G_t, s_{ik} is unit i's exposure to $k \in \{G, R\}$ and γ measures the differential response to R, conditional on s_{ir}. β_h^P is the portable elasticity at horizon h.

This expression highlights that two objects determine whether the endogenous response of R to ε_{gt} is fully differenced out. First, a cross-sectional term—denoted by ϕ—that measures whether exposure to G is correlated with exposure to R. For example, are regions that receive more military spending also more or less sensitive to interest rate changes? The cross-sectional

[3] Concretely, consider a unit-level outcome that evolves according to: $Y_{it} = \beta \times s_{ig} G_t + \gamma \times s_{ir} R_t + u_t + u_i + u_{it}$.

term will be different from zero when the sensitivities of economic units to G and R are correlated in the cross-section.

Second, a time-series object—denoted by $\boldsymbol{R}_{g,h}$—that captures the impulse response of R to the shock of interest, ε_{gt}, at horizon h. For example, in a fiscal setting, this could represent the response of monetary policy to government spending or tax policy shocks. The time series term will be different from zero when the shock of interest triggers an endogenous general equilibrium (GE) response in R.

The TWFE specification implicitly assumes that the cross-sectional heterogeneity used for identification only drives exposure to the shock of interest and not to any aggregate variable that the researcher intends to difference out. This is a strong assumption for many empirical applications where exposure shifters often reflect structural characteristics of economic agents (e.g., demographics, industrial composition, financial constraints). It is plausible that such characteristics drive the response to multiple aggregate variables.

This poses a clear identification challenge: we cannot rely on time-fixed effects alone to control for general equilibrium responses. The estimated coefficients reflect not only the direct effects of the shock of interest but also the effects through the endogenous response in other aggregate variables:

$$\hat{\beta}_h \rightarrow \beta_h^P + \underbrace{\gamma \times \phi \times \boldsymbol{R}_{g,h}}_{\Omega_h}. \tag{1.3}$$

This second component, which I label the HEDGE term and denote by Ω_h, acts as an additional treatment effect. It measures how the endogenous and dynamic response of R to ε_{gt} affects the outcome of units that are relatively more exposed to G. Specifically, for a given general equilibrium response $\boldsymbol{R}_{g,h}$, the contribution of the HEDGE term depends on the covariance between exposure to the shock of interest and exposure to the general equilibrium response.

This creates two concerns. First, the sign of the bias introduced by HEDGE is not dictated by economic theory alone, but by the sign of ϕ, the scaled covariance between exposure shifters in the cross-section. As a result, it is possible for general equilibrium forces that attenuate the aggregate effect of a shock, such as contractionary monetary policy in response to a fiscal expansion, to increase the estimated cross-sectional effect, depending on the structure of the data. Second, exposure shifters used in different studies can be correlated with sensitivities to different aggregate variables, leading to variation in estimated effects even when studying the same aggregate shock. These issues highlight the importance, both for interpretation and for empirical consistency, of separating the direct effect from the component driven by HEDGE.

Decomposition Framework

I develop a framework that brings in additional information to identify portable elasticities in HEDGE economies. Its goal is to decompose $\hat{\beta}_h$ in expression (1.3) into a portable component β_h^P and an HEDGE component Ω_h. Specifically, I develop an estimation procedure for the latter term. The framework is based on the results in McKay and Wolf (2023) and Barnichon and Mesters (2023), who show that reduced-form impulse responses to contemporaneous and news shocks are sufficient to construct dynamic counterfactuals in a general class of linear models. I adapt these results to a different object of interest: the identification of cross-sectional elasticities clean of HEDGE.

The methodology combines cross-sectional and time-series analysis in three steps. First, a cross-sectional step estimates the response of the outcome Y_{it} to an innovation in R, the aggregate variable we intend to difference out, conditional on s_{ig}, the exposure shifters used for identification. Second, a time series step estimates a sequence of counterfactual innovations to R that replicate the impulse response of R to the shock of interest ε_{gt}. The final step uses the cross-sectional responses from the first step to evaluate the propagation of the innovations estimated in the second step. This provides an estimate of Ω_h for each horizon h. The estimated portable elasticity at horizon h, $\hat{\beta}_h^P$, is then given by $\hat{\beta}_h - \hat{\Omega}_h$.

Applying the decomposition requires one additional input relative to the TWFE strategy:

exogenous variation in the aggregate variable (or variables) that confound identification. The specific nature of this variation — whether contemporaneous or involving news (i.e., anticipated) shocks — depends on the structure of the data-generating process. I distinguish three cases: static, dynamic without anticipation effects, and dynamic with anticipation effects, each imposing different informational requirements for identifying portable elasticities.

Before turning to the implementation of the decomposition, I show that the same source of exogenous variation can also be used to test the presence of HEDGE. The test is straightforward to implement within the TWFE specification: it requires adding an interaction between s_{ig} and the source of exogenous variation in the aggregate variable of interest. For example, one can test whether exposure to fiscal policy is correlated with exposure to monetary policy by adding an interaction between s_{ig} and the interest rate, instrumenting the latter with a well-identified series of monetary shocks. A statistically significant coefficient on this interaction term is evidence of HEDGE, provided the aggregate variable responds to the shock of interest. Importantly, the test relies only on s_{ig}, the source of cross-sectional variation being used for identification, and does not require knowledge of the true exposure shifters for other aggregate variables, such as s_{ir} in the above example. If the test finds evidence of HEDGE, the decomposition framework uses the same exogenous variation to estimate and remove the associated bias.

In static environments, fluctuations are driven by purely transitory shocks, and the framework can be implemented using any observed contemporaneous shock to the GE variable R. Although of little empirical relevance, I present this static setting to discuss how my framework relates to a common robustness check used in the cross-sectional literature. This robustness check, often referred to as *control function approach*, aims to address the HEDGE omitted variable bias by including an interaction term between the exposure shifter to the shock of interest, s_{ig}, and the endogenous aggregate variable R that the researcher suspects is not being differenced out. The underlying assumption is that this interaction term controls for the differential effects of R, without the need to use exogenous variation in R. In static settings, I show that my framework and the control function approach yield equivalent results. They both identify the true portable elasticity.

However, this equivalence result breaks down in dynamic settings, a characteristic feature of most empirical applications. Once dynamics are at play, the control function approach yields inconsistent estimates of portable elasticities because it cannot control for the *dynamic* path of the GE variable *conditional* on the shock of interest. It can only control for dynamic paths of the GE variable that lie in the subspace generated by variation in the included controls. In the simplest case with only a control for R_t, this corresponds to the unconditional autocovariance of R_t.[4] By contrast, the decomposition framework provides an estimation strategy that explicitly accounts for the dynamic behavior of macroeconomic variables. Thus, my approach is robust to general macroeconomic settings and represents an improvement over the control function approach.

The dynamic settings can be separated into two cases. First, I show how to implement the framework in dynamic settings without anticipation or forward-looking effects—referred to as Markov economies. In such cases, identification requires no further information beyond the static case: contemporaneous innovations in R suffice to identify the HEDGE term and the portable elasticity across horizons.

In settings with forward-looking dynamics, where both contemporaneous and expected realizations of variables affect current responses, identification of Ω_h additionally requires news shocks to the expected future path of the GE variable (Barnichon and Mesters, 2023; McKay and Wolf, 2023). Because such shocks are difficult to obtain, I outline how to apply the framework under limited information and then employ model simulations to assess estimation performance. Across both Markov and forward-looking economies, the framework improves upon the TWFE and control function approaches, even when only contemporaneous innovations to R_t are available.

[4]Intuitively, the impulse response of R depends on the shock that drives it. For example, monetary policy may respond to a demand shock with a hump-shaped path, while orthogonal innovations in R may follow a more AR(1)-like process. A single contemporaneous control cannot distinguish between these shock-contingent responses.

Empirical Application

I illustrate how to implement the framework by estimating US cross-sectional fiscal multipliers using state-level data, perhaps the most influential application of cross-sectional methods in macroeconomics. There is consensus that an advantage of cross-sectionally identified multipliers is their independence from GE effects, in particular those operating through the response of monetary policy. I study whether this is the case: Is monetary policy differenced out in the cross-section of US states? For estimation, I use the dataset on local defense spending built by Dupor and Guerrero (2017) and the series of Romer and Romer (2004) monetary shocks. Using specification (1.1), which is standard in the literature, I estimate a cumulative 2-year multiplier of 1.5. Applying the decomposition framework, I estimate a *portable* multiplier of 1. This estimate should be interpreted as a *monetary-policy–portable multiplier*, that is the cross-sectional effect of government spending holding the monetary response fixed.[5]

The difference with the TWFE estimate implies a positive HEDGE term. This arises because US states that are more exposed to defense spending shocks are also less sensitive to interest rates and, on average, interest rates increase in response to defense spending shocks. Together, these two imply an upward bias in the TWFE estimate.[6] This finding challenges the assumption that TWFE estimates of the cross-sectional fiscal multiplier are independent from the monetary regime and underscores the importance of GE effects in shaping cross-sectional estimates.

To interpret the empirical estimates, I build a two-region New Keynesian model with heterogeneity in exposure to fiscal and monetary policy. The framework combines the structure of Nakamura and Steinsson (2014), who study government spending in a two-region representative

---

[5]While the framework accommodates multiple confounding GE variables, I leave the analysis of additional sources of GE feedback to future work. Importantly, as discussed below, removing this monetary policy channel already restores a central use of cross-sectional multipliers: it allows researchers to combine the estimated effect with alternative monetary policy regimes and still make valid inferences on the aggregate effects of fiscal policy.

[6]These findings can be translated into the notation used to define Ω_h as follows. Increases in the interest rate have contractionary effects on local production, so $\gamma < 0$. The fact that high s_{ig} states are less sensitive to interest rates implies $\phi < 0$ and a monetary tightening following the fiscal shock implies $\boldsymbol{R}_{g,h} > 0$. Taken together, this yields $\Omega_h = \gamma \times \phi \times \boldsymbol{R}_{g,h} > 0$ and an upward bias in the TWFE estimates.

agent setting, with the two-region TANK features of Herreño and Pedemonte (2025), who introduce hand-to-mouth households but abstract from fiscal shocks. The model delivers theoretical analogs of the TWFE cross-sectional multiplier, the HEDGE term, and the portable multiplier and allows me to decompose the cross-sectional effect into distinct channels. In particular, I separate the usual demand forces—expenditure switching across regions and intertemporal substitution driven by relative price changes—from a HEDGE channel that captures intertemporal substitution driven by changes in the common national rate. This channel arises when regions differ in their interest rate sensitivity and acts as an additional treatment that pushes the TWFE cross-sectional multiplier away from its portable value.

Previous work has argued that cross-sectional multipliers can be interpreted as bounds on the aggregate effects of fiscal policy, serving as an upper bound under conventional monetary policy, where contractionary interest rate responses are differenced out, and as a lower bound at the zero lower bound (Chodorow-Reich, 2019; Dupor et al., 2023). This bounding logic relies on the cross-sectional design cleanly separating the direct effect from the aggregate *missing intercept* (Wolf, 2023b) operating through monetary policy. With HEDGE, this separation fails: time fixed effects no longer fully remove the monetary response. As a result, the TWFE cross-sectional multiplier can vary widely across monetary regimes and need not bound the aggregate effect. In contrast, the portable multiplier remains stable across monetary regimes and retains its interpretation as a valid bound.

Finally, the model delivers a sufficient statistic that summarizes the precision loss from relying on limited news shocks. Calibrating this statistic to the US setting indicates that the decomposition recovers the portable multiplier with only limited error, even in a forward-looking New Keynesian setting.

Related Literature

This paper contributes to three strands of the macroeconomics literature. First, it is related to the empirical literature that uses cross-sectional methods to identify the effects of aggregate shocks. Nakamura and Steinsson (2018) provide a general review of the state and scope of this literature, while Chodorow-Reich (2020) focuses on the specifics of cross-regional analysis. Nakamura and Steinsson (2018) introduce the concept of portable statistics as well-identified moments or parameters that can be used to distinguish between models and imported across economic settings, precisely due to their partial equilibrium interpretation.

The typical empirical design in this literature relies on time fixed effects[7] to absorb GE effects. Papers that recognize the limitations of time fixed effects usually address the problem through a control-function approach, incorporating an interaction between the exposure shifter used for identification and the aggregate variable that varies in GE. First, I show that, in typical macroeconomic settings, the control function approach generally fails to control for HEDGE. Second, I propose a methodological framework that identifies PE elasticities in HEDGE economies and is robust to the dynamic patterns of macroeconomic data. The estimation framework in this paper extends the use of cross-sectional methods to macroeconomic settings in which cross-sectional variation plus time fixed effects are insufficient to control for dynamic GE effects.

This paper complements a recent literature that studies other limitations of cross-sectional designs to estimate the effects of aggregate shocks. Canova (2024) proposes an estimator robust to heterogeneity in the autoregressive dynamics of local outcomes. Adao et al. (2019a) and Manuel (2025) show how to incorporate spatial links between regions, which allows indirect shock spillovers (through trade and migration) to be accounted for in local impact estimates. On a different front, Almuzara and Sancibrián (2024) focus on inference in these settings, advancing a procedure to obtain robust standard errors with correct coverage.

The study closest to this paper is Arkhangelsky and Korovkin (2024), who propose a reweighting procedure to identify contemporaneous responses to aggregate shocks in the presence of unobserved aggregate confounders. Their synthetic control approach constructs weights that orthogonalize exposure to the shock of interest from exposure to these confounders in the cross

[7]With a single cross-section, the intercept plays the role of time fixed effects in panel settings.

section.

This paper is complementary and takes a different route to the same identification challenge. I focus on environments in which aggregate variables are observed but respond dynamically to the shock. Rather than reweighting the cross section, I use time-series variation to construct counterfactual shocks that offset the GE response of observed aggregate variables.[8] This approach accommodates settings in which the same exposure drives sensitivity to multiple aggregates and where no single reweighting scheme can eliminate all sources of bias.

The decomposition framework that I propose is also closely related to the re-centered instrument in Borusyak and Hull (2023). In their setting, the shift-share instrument is re-centered by subtracting its expected value across possible shock assignments, so that it isolates variation arising from higher-than-expected exposure to exogenous shocks. My approach can be thought of as re-centering the aggregate shock of interest by removing the expected GE response that it induces in other aggregates.

The question of how to translate cross-sectional effects into aggregate effects has led to a growing literature on aggregation. Aggregation strategies include back-of-the-envelope calculations, model-based aggregation, and econometric frameworks that employ panel and factor structures (Sarto, 2024; Matthes et al., 2024)[9]. Several papers combine reduced-form estimates at different levels of aggregation with structural assumptions, as in Guren et al. (2021b), Chodorow-Reich et al. (2021), Guren et al. (2021a), Wolf (2023b) and Wolf (2019). I complement this strand of work with a method to identify the free-of-GE micro elasticities that serve as starting points for both back-of-the-envelope and model-based aggregation.

The methodological framework that I propose is closely related to recent advances in the

[8]To fix ideas, consider a setting with a single aggregate confounder R. The standard identifying condition requires $cov(s_{ig}\varepsilon_{gt}, s_{ir}R_t) = 0$, which fails in the presence of HEDGE. My approach can instead be interpreted as constructing a *counterfactual instrument* $\tilde{z}_{it} = s_{ig}(\varepsilon_{gt} - \tilde{\varepsilon}_{rt})$, where $\tilde{\varepsilon}_{rt}$ denotes the counterfactual shock that offsets the GE response of R_t to ε_{gt}. This implies that $cov(\tilde{z}_{it}, s_{ir}R_t) = 0$, even when no reweighting of the cross section can achieve this orthogonality. This construction illustrates how time-series information can be used to undo the GE response by orthogonalizing shocks, rather than shares.

[9]These papers propose econometric frameworks that estimate aggregate effects directly using a Bayesian approach in Matthes et al. (2024) and a factor model in Sarto (2024).

computation of macroeconomic counterfactuals made by McKay and Wolf (2023), Barnichon and Mesters (2023) and Caravello et al. (2024). McKay and Wolf (2023) show that reduced-form time series responses to contemporaneous and news policy shocks are sufficient to construct policy rule counterfactuals in a general class of linearized models. Barnichon and Mesters (2023) further show that these impulse responses are also enough to characterize the optimality of the policy rule. My decomposition framework builds on their work and extends it to the cross-sectional setting. I show how to combine cross-sectional and time series reduced-form responses to aggregate shocks to purge cross-sectional elasticities from HEDGE. The decomposition that I propose can be interpreted as computing cross-sectional elasticities under the counterfactual scenario of homogeneous exposure to GE.

The empirical application speaks to the literature on cross-sectional fiscal multipliers. Examples of this literature are Nakamura and Steinsson (2014), Chodorow-Reich et al. (2012), Auerbach et al. (2020), Demyanyk et al. (2019), Dupor and Guerrero (2017), Dupor et al. (2023), Pennings (2022), and Pennings (2021), among others. Chodorow-Reich (2019) provides a comprehensive review of the lessons from this literature. One of the advantages attributed to cross-sectional methods for the estimation of fiscal multipliers is their independence from the state of the economy. In particular, their independence from the stance of monetary policy.[10]

At the same time, it is well documented that sensitivities to interest rates vary substantially between economic units (e.g., households, firms, or regions). For example, Herreño and Pedemonte (2025) find that US cities with lower per capita income are more responsive to monetary policy, a pattern that they show is consistent with a New Keynesian model of a monetary union with heterogeneous shares of hand-to-mouth households. Building on these insights, I revisit the estimation of cross-sectional fiscal multipliers by explicitly accounting

[10]The strength of the monetary policy response to fiscal shocks is an important determinant of the size of the aggregate fiscal multiplier. Theoretical and empirical papers on the interaction between fiscal and monetary policy include Christiano et al. (2011), Canova and Pappa (2011), DeLong and Summers (2012), Bachmann and Sims (2012), Riera-Crichton et al. (2015), Auerbach and Gorodnichenko (2015), Farhi and Werning (2016), Leeper et al. (2017), Kato et al. (2018), Ramey and Zubairy (2018), Ramey (2019), Cloyne et al. (2020), Hack et al. (2023), among many others.

for heterogeneity in monetary policy responses. My contribution is twofold. First, I show that heterogeneity in monetary policy responses across US states is correlated with their exposure to defense shocks. This implies that time fixed effects do not *fully* difference out monetary policy responses. Second, I introduce new estimates that control for these differential effects, thereby providing estimates of the cross-sectional fiscal multiplier that can be more confidently interpreted as independent of the monetary policy stance.

Layout

The remainder of the paper is structured as follows. Section 1.2 discusses the identification challenge at the core of the paper and presents a diagnostic test for it. Section 1.3 presents the estimation framework. Section 1.4 presents the empirical application. Section 1.5 presents a 2-region New Keynesian model to interpret the empirical findings from the previous section. Lastly, Section 1.6 concludes.

1.2 When Does HEDGE Challenge Portability?

First, this section introduces a simple framework to illustrate the identification challenge that arises when units are differentially exposed to aggregate variables that co-move with the shock of interest. Next, it presents a diagnostic test to detect whether HEDGE is present with respect to some aggregate variable R that the researcher intends to difference out.

To build intuition, consider a simple HEDGE economy in which units differ in their exposure to an aggregate shock of interest ε_{gt} and to one other aggregate variable R_t:

$$Y_{it} = \beta s_{ig}\varepsilon_{gt} + \gamma s_{ir}R_t + u_i + u_t + u_{it}, \tag{1.4}$$

where Y_{it} is the outcome (e.g., local GDP) of unit i in period t and s_{ik} denotes the unit-specific exposure shifter to the aggregate variable k. Unit, time, and unit-time idiosyncratic shocks are denoted with a u plus the corresponding sub-indices. It can be shown that specification (1.1)[11],

[11]In terms of this example, specification (1.1) is $Y_{it+h} = \beta_h s_{ig}\varepsilon_{gt} + \lambda_{ih} + \lambda_{th} + e_{it+h}$.

which uses unit and time fixed effects, recovers:

$$\hat{\beta}_h \rightarrow \beta_h^P + \gamma \underbrace{\frac{cov\left[s_{ig}, s_{ir}\right]}{v[s_{ig}]}}_{\phi} \underbrace{\frac{cov\left[R_{t+h}, \varepsilon_{gt}\right]}{v[\varepsilon_{gt}]}}_{\boldsymbol{R}_{g,h}}. \tag{1.5}$$

The above expression clarifies the conditions under which the dynamic effects that operate through R are not fully differenced-out. There are two distinct objects: one related to time series variation, $\boldsymbol{R}_g = \{\boldsymbol{R}_{g,h}\}_h$, and one related to the source of cross-sectional variation used for identification, ϕ. The former measures the impulse response of R to the G shock of interest. In other words, $\boldsymbol{R}_g$ captures the dynamic path of R conditional on the realization of the shock of interest. For example, in a fiscal application, this could be the response of federal tax rates to federal government spending shocks.

The cross-sectional term, ϕ, measures the degree to which unit-specific exposure shifters to different aggregate variables co-move with each other. It informs whether the driver of exposure to the shock of interest ε_{gt} is systematically related to the driver of exposure to R. For example, are regions that receive more government spending systematically more or less sensitive to taxation or to monetary policy? Expression (1.5) highlights that the assumption in a TWFE strategy like (1.1) is that the source of cross-sectional heterogeneity used for identification only drives variation in exposure to the shock of interest and not to any of the $\{R_t^k\}_k$ variables that the researcher intends to difference out:

$$cov[s_{ig}, s_{ik}] = 0 \quad \forall\, k \text{ with } cov[R_{t+h}^k, \varepsilon_{gt}] \neq 0 \text{ for some } h. \tag{1.6}$$

This assumption may be hard to satisfy in several macroeconomic applications where the type of unit-level characteristics that we use to proxy for heterogeneity in exposure to one aggregate shock are also likely to affect exposure to other aggregate variables. For example, in the fiscal literature, constrained or hand-to-mouth households are an important determinant of the sensitivity of output to fiscal stimulus. These households play a similar role when it comes to the sensitivity of output to monetary policy shocks.

Violations of the identifying assumption lead to omitted-variable bias, whose sign and magnitude can vary depending on the source of heterogeneity. As a result, studies using different sources of cross-sectional variation (or focusing on different time periods) may estimate different elasticities for the same aggregate shock, due to loading on different aggregate variables that move in GE. Such situations decrease the portability of cross-sectionally identified elasticities not just because some of the general-equilibrium responses are part of the estimates, but also because how these GE responses affect estimates depends on the cross-sectional and time-series variation that is being used.

1.2.1 Test for HEDGE

Next, I present a test to diagnose whether HEDGE poses an identification threat. In economic terms, two conditions must be satisfied for HEDGE with respect to a given aggregate variable R to pose an identification problem. First, R must respond to the shock of interest; that is, it must be one of the macroeconomic variables that move in general equilibrium. Second, at the cross-sectional level, units with different exposures to the shock of interest must exhibit heterogeneous responses to changes in R.

I assume that the researcher has identified a set of aggregate variables $\{R_t^k\}_{k\in K}$ that respond in GE to the shock of interest, in the sense that $cov(R_t^k, \varepsilon_{gt}) \neq 0$, and that she intends to difference out. This set can be informed by time-series econometric tools, economic theory, or a combination of both. The test proposed here focuses on the second condition for HEDGE, namely whether the cross-sectional variation used for identification is orthogonal to differential responses to these equilibrium variables.

The goal is then to test whether s_{ig} drives heterogeneity in cross-sectional responses to a given aggregate variable R. Let ε_{rt} denote a source of exogenous variation in R (e.g., a monetary policy shock). The test exploits variation in ε_{rt} to identify differential responses to R, conditional on s_{ig}. It can be implemented via the following regression:

$$Y_{it+h} = c_h \times s_{ig} R_t + \lambda_{ih} + \lambda_{th} + e_{it+h} \qquad \text{for all} \quad h \in H, \tag{1.7}$$

where the interaction $s_{ig}R_t$ is instrumented with $s_{ig}\varepsilon_{rt}$, and λ_{ih}, λ_{th} are horizon-specific unit and time fixed effects. In regression (1.7), c_h captures how the response of Y_{it+h} to changes in R_t varies with s_{ig}. For aggregate variables that respond to the shock of interest, a statistically significant coefficient indicates that units more exposed to G also respond differently to changes in R at horizon h, violating the identification assumption of the TWFE approach. In sum, evaluating HEDGE in relation to an aggregate variable R amounts to testing whether these coefficients are jointly zero:

$$H_0: \ c_h = 0 \ \forall \ h \in H, \qquad\qquad H_A: \ c_h \neq 0 \ \text{ for some } \ h \in H. \tag{1.8}$$

Rejection of the null hypothesis implies that units with different levels of exposure to the shock of interest, as captured by s_{ig}, exhibit systematically different responses to changes in the aggregate variable R. In this case, estimates from cross-sectional regressions that rely on time fixed effects are likely to confound the direct effect of the shock with general equilibrium responses operating through R. Detecting such patterns in the data is a signal that HEDGE may be an identification threat and that a correction is necessary, such as the decomposition framework proposed in Section 1.3.

The test can be easily generalized to assess exposure heterogeneity with respect to multiple aggregate variables. Specifically, several interaction terms of the form $s_{ig}R_t^k$, each instrumented with $s_{ig}\varepsilon_{rt}^k$, can be included for a set of variables $\{R^k\}_{k\in K}$. A joint test of significance on the corresponding coefficients $\{\hat{c}_h^k\}$ then provides a diagnostic for which general equilibrium channels pose a threat to identification in the cross section.

Importantly, the informational cost of this diagnostic is relatively low. The test requires only a source of exogenous variation in the aggregate variable R that the researcher aims to difference out. This variation is the same as that used to study the effects of R itself, for example, monetary or tax policy shocks, and can often be recovered using structural time series methods such as VARs or narrative identification approaches. In addition, the test does not require knowledge of the true driver of unit-level exposure to R. This is because, in a linear regression framework, what matters for the estimation is the linear relationship between the true exposure shifter and s_{ig}. Therefore, as long as s_{ig} is used for identification and is linearly related to the true exposure shifter, the test remains valid.

Recap

This section clarifies when HEDGE poses a threat to the identification of portable elasticities. Specifically, HEDGE threatens the validity of the TWFE strategy when (i) the GE variable responds to the shock of interest and (ii) the exposure shifter used for identification is correlated with heterogeneity in responses to that GE variable. I provide a simple test to detect such threats by leveraging exogenous variation in the GE variable. A rejection of the null signals that the identifying variation used in the cross-section is also picking up differential responses to general equilibrium forces. In such cases, cross-sectional estimates are likely to conflate the direct effect of the shock with these indirect channels. The next section presents a decomposition framework to estimate portable elasticities by separating them from the HEDGE component.

1.3 Decomposition Framework

This section presents a decomposition framework to identify portable elasticities in HEDGE economies. The goal is to decompose $\hat{\beta}_h$, the elasticity estimate for horizon h of the TWFE specification, into two distinct components: a portable component, β^P_h, and a HEDGE component, denoted by Ω_h.

The HEDGE component Ω_h measures how the response of a GE variable, namely R, affects the outcome of units with different levels of exposure to the shock of interest. In summary, the framework I propose uses reduced form cross-sectional responses to R to remove the effect of the expected response of R from the TWFE estimates. I outline how to implement this procedure under varying assumptions about the structure of the economic environment. In each case, identifying the HEDGE term requires access to exogenous variation in R, but the nature of this variation depends on the specific setting. I distinguish three stylized cases in terms of their informational requirements and describe how the methodology applies to each.

The first case corresponds to static economic environments. Here, both aggregate and cross-sectional variables are assumed to be driven by purely transitory, serially uncorrelated shocks. Implementing the framework requires access to contemporaneous innovations in the GE variable, R. Although this case has limited empirical relevance for macroeconomic questions, it provides a useful benchmark to compare the proposed framework with the control function approach, which consists of including the interaction $s_{ig}R_t$ as a control in the estimating equation. I show that under the static setting assumptions, the two approaches yield equivalent results. Both consistently estimate portable elasticities. However, the control function approach yields inconsistent estimates when dynamics are introduced, whereas the proposed framework remains valid.

The second case considers dynamic environments that admit a finite-order vector autoregression (VAR) representation (i.e. environments that satisfy a Markov structure). In this environment, implementing the framework requires contemporaneous innovations in R, just as in the static setting. This allows identification of the HEDGE term, Ω_h, both at impact and at different horizons ($h > 0$).

The third and most general case extends the framework to forward-looking environments, where both current and expected future realizations of variables affect present outcomes, such as in DSGE models. In these settings, identifying the HEDGE terms requires information not only on contemporaneous innovations in R, but also on news shocks that shift expectations about the future path of R. This extension is built on the work of McKay and Wolf (2023) and Barnichon and Mesters (2023), who show how time-series reduced-form responses to contemporaneous and news shocks can be used to construct counterfactuals at the aggregate level.

The remainder of this section presents each case in detail. Then, it briefly discusses the argument to establish consistency and how to conduct inference. Detailed explanations are relegated to Appendix A.1.6 and Appendix A.1.7, respectively.

1.3.1 Static Setting

Consider an economy that can be characterized by the following system of equations:

$$
\begin{aligned}
G_t &= \alpha R_t + u_{gt} \qquad & Y_{it} &= \beta s_{ig} G_t + \gamma s_{ir} R_t + u_i + u_t + u_{it}^{y} \\
R_t &= \delta G_t + u_{rt} & s_{ig} &\sim N(\bar{s}, \sigma_{s_g}^2) \\
u_{kt} &\sim N(0, \sigma_k^2) & s_{ir} &= \phi s_{ig} + u_i^{s_r}, \quad s_{ig} \perp u_i^{s_r} \\
& & u_i^{s_r} &\sim N(0, \sigma_{s_r}^2), \quad u_{it}^{y} \sim N(0, \sigma_y^2)
\end{aligned}
\tag{S1}
$$

where i denotes economic units (e.g., regions, households, firms) and t time. G_t and R_t are two aggregate variables that are allowed to endogenously respond to each other. u_{gt} and u_{rt} have the interpretation of structural innovations to G and R, respectively. The unit-level outcome Y_{it} is a function of both G_t and R_t plus unit, time and unit-time idiosyncratic innovations. The cross-sectional variation in exposure to G and R is governed by s_{ig} and s_{ir}, respectively. I assume that s_{ir}, the exposure shifter to R_t, is linearly related to s_{ig}, the exposure shifter to G_t, with ϕ governing the strength of this relation. When $\phi = 0$, the exposure shifters to G and R are uncorrelated in the cross section. I further assume that the cross-sectional exposure shifters, s_{ig} and s_{ir}, are time-invariant and, therefore, independent of aggregate shocks.

The economy is static in the sense that cross-sectional and aggregate variables respond only contemporaneously to shocks, with no propagation over time. Given the absence of dynamics, the portable elasticity is non-zero only at impact

$$\beta_h^P = \begin{cases} \beta & \text{if } h = 0 \\ 0 & \text{if } h > 0 \end{cases} \tag{1.9}$$

and measures the relative effect of a change in G on the unit-level outcome, holding R fixed. The static nature of the problem implies that the cross-sectional elasticity for any horizon $h > 0$ is equal to zero, thus I drop the subscript h for the remainder of this subsection.

For illustration purposes, I consider a simplified setting with $\alpha = 0$, shutting down the endogenous response of G_t to R_t. The estimation steps and results that follow equally apply to a setting with two-way feedback as shown in Appendix A.1.1. I also normalize the variance of idiosyncratic shocks and the shares s_{ig} to unity.

Consider a researcher who observes an innovation ε_{gt} to G_t satisfying Assumption 1.1.

Assumption 1.1. *The observed innovation ε_{gt} is a noisy measure of the structural shock u_{gt}:*

$$\varepsilon_{gt} = \kappa u_{gt} + \nu_{gt}, \quad \kappa \neq 0, \tag{1.10}$$

where $\nu_{gt} \sim N(0, \sigma^2_{\nu_g})$ is orthogonal to all structural innovations, in particular

$$cov[\nu_{gt}, u_{gt}] = 0, \quad cov[\nu_{gt}, u_{rt}] = 0. \tag{1.11}$$

This assumption ensures that ε_{gt} isolates variation in G_t orthogonal to the structural innovation in R_t. More generally, ε_{gt} may load on other aggregate innovations, in which case the decomposition that follows purges R-driven effects but may still reflect transmission through those channels.[12]

[12] See Appendix A.1.3 for details.

The typical estimation strategy consists of the following TWFE regression:

$$Y_{it} = b \times s_{ig} G_t + \lambda_i + \lambda_t + e_{it}. \tag{1.12}$$

Because G_t may be endogenous, $s_{ig}\varepsilon_{gt}$ is used as an instrument for $s_{ig}G_t$. Here, λ_i and λ_t denote unit and time fixed effects, respectively. For ease of exposition, I set κ such that $cov[G_t, \varepsilon_{gt}]/v[\varepsilon_{gt}] = 1$. The coefficient then converges in probability to the following expression:[13]

$$\begin{aligned} \mathbb{E}[\hat{b}] - \beta^P &= \gamma \times \frac{cov[s_{ir}R_t, s_{ig}\varepsilon_{gt}]}{cov[G_t, \varepsilon_{gt}]} \\ &= \gamma \times \phi \times \frac{cov[R_t, \varepsilon_{gt}]}{cov[G_t, \varepsilon_{gt}]} \end{aligned} \tag{1.13}$$

Here γ is the structural coefficient governing the response of Y_i to a change in R_t, conditional on exposure s_{ir}, and ϕ captures the cross-sectional correlation between s_{ig} and s_{ir}. The product $\tilde{\gamma} = \gamma \times \phi$ therefore captures the response of Y_i to a change in R_t, conditional on s_{ig} rather than s_{ir}:[14]

$$\mathbb{E}[\hat{b}] - \beta^P = \tilde{\gamma} \times \frac{cov[R_t, \varepsilon_{gt}]}{v[\varepsilon_{gt}]} = \Omega, \tag{1.15}$$

where Ω denotes the bias associated with HEDGE. In order to decompose $\hat{b}$ into a portable and an HEDGE component, I propose an estimator for Ω, such that :

$$\mathbb{E}[\hat{\beta}^P] = \mathbb{E}[\hat{b} - \hat{\Omega}] = \beta^P. \tag{1.16}$$

This is the key idea of the framework. The additional piece of information required to estimate Ω

[13] Throughout, expectations of estimated coefficients refer to their probability limits.

[14] Replacing s_{ir} for its expression in terms of ϕ, s_{ig} and $u_i^{s_r}$ in the unit-level outcome yields:

$$\begin{aligned} Y_{it} &= \beta \times s_{ig} G_t + \gamma \times \phi \times s_{ig} R_t + \gamma \times u_i^{s_r} R_t + u_i + u_t + u_{it}^y \\ &= \beta \times s_{ig} G_t + \tilde{\gamma} \times s_{ig} R_t + \gamma \times u_i^{s_r} R_t + u_i + u_t + u_{it}^y, \end{aligned} \tag{1.14}$$

where $\tilde{\gamma}$ is the parameter associated to the interaction term between s_{ig} and R_t.

is an instrument ε_{rt} for R_t that satisfies the following:

Assumption 1.2. *The observed innovation ε_{rt} is a noisy measure of the structural shock u_{rt}:*

$$\varepsilon_{rt} = \upsilon u_{rt} + \eta_{rt}, \quad \upsilon \neq 0,$$

where $\eta_{rt} \sim N(0, \sigma^2_{\eta_r})$ is orthogonal to structural innovations, in particular:

$$cov[\eta_{rt}, u_{gt}] = 0, \quad cov[\eta_{rt}, u_{rt}] = 0. \tag{1.17}$$

Unlike ε_{gt} , the orthogonality restriction on η_{rt} cannot be relaxed. Any loading of η_{rt} on other structural innovations would induce correlation between ε_{rt} and these shocks, violating the exclusion restriction necessary for instrument validity.

Equipped with ε_{gt} and ε_{rt}, the estimation proceeds in three steps. For clarity, I describe the procedure in terms of population moments; the empirical implementation replaces these with their sample counterparts. Throughout, I denote the contemporaneous response of X to the observed innovation ε_{kt} by $\boldsymbol{X}_{k,t} = cov[X_t, \varepsilon_{kt}]/v[\varepsilon_{kt}]$.

1. *Cross-sectional Step.* Estimate the coefficients from a regression of Y_{it} on $s_{ig}\varepsilon_{gt}$ and $s_{ig}\varepsilon_{rt}$. These coefficients converge in probability to:

$$\beta + \tilde{\gamma} \times \boldsymbol{R}_{g,t}, \qquad \tilde{\gamma} \times \boldsymbol{R}_{r,t},$$

   respectively.

2. *Time Series Step.* The response of R_t to a unit change in ε_{rt}, given by $\boldsymbol{R}_{r,t}$, may differ from its response to a unit change in ε_{gt}, given by $\boldsymbol{R}_{g,t}$. This step estimates these two time-series impulse responses and then rescales ε_{rt} to match the response of R_t to ε_{gt}:

$$\boldsymbol{R}_{r,t} \times \tilde{\varepsilon}_{rt} = \boldsymbol{R}_{g,t} \Rightarrow \tilde{\varepsilon}_{rt} = \frac{\boldsymbol{R}_{g,t}}{\boldsymbol{R}_{r,t}}.$$

3. *Final Step.* Combine the cross-sectional response to ε_{rt} with the rescaling from the time-series step to obtain:

$$\overbrace{\tilde{\gamma} \times \boldsymbol{R}_{r,t}}^{\text{CS Step}} \times \underbrace{\frac{\boldsymbol{R}_{g,t}}{\boldsymbol{R}_{r,t}}}_{\text{TS Step}} = \Omega.$$

   This constructs the counterfactual response of Y_{it} to the component in R_t induced by ε_{gt}, isolating the HEDGE contribution to the baseline estimate.

Next, I present a detailed discussion on the implementation of each step.

Cross-sectional Step

The goal of this step is to identify the reduced form response of Y_i to both G and R, conditional on exposure to the shock of interest s_{ig}. As outlined above, this requires access to an instrument for R_t. For example, if R_t is the interest rate, then a series of well-identified monetary policy shocks would satisfy this requirement. The reduced form responses to changes in G_t and R_t can be estimated as follows:

$$Y_{it} = b \times s_{ig}\varepsilon_{gt} + c \times s_{ig}\varepsilon_{rt} + \lambda_i + \lambda_t + e_{it}, \tag{1.18}$$

where each aggregate shock is interacted with s_{ig}, the exposure shifter to G_t. It can be shown that this specification recovers the following two coefficients:

$$\begin{aligned} \mathbb{E}[\hat{b}] &= \beta^P + \tilde{\gamma} \times \boldsymbol{R}_{g,t}, \qquad\qquad \mathbb{E}[\hat{c}] = \tilde{\gamma} \times \boldsymbol{R}_{r,t}, \\ &= \beta^P + \underbrace{\tilde{\gamma} \times \delta}_{\Omega} \end{aligned} \tag{1.19}$$

where δ captures the response of R to an innovation in G. First, the coefficient on $s_{ig}\varepsilon_{rt}$ isolates variation in R_t that is orthogonal to G_t, and therefore identifies $\tilde{\gamma}$. Second, because this regression only exploits the variation in R_t coming from ε_{rt}, the coefficient $\hat{b}$ is still a combination of the portable and HEDGE components as in the TWFE case. In this reduced-form specification, we isolate the variation in R_t driven by ε_{rt}, explicitly omit the endogenous variation in R_t induced by G_t.

Alternatively, one could augment the TWFE regression (1.12) by incorporating $s_{ig}R_t$ as an additional interaction:

$$Y_{it} = b^{CF} \times s_{ig}G_t + c^{CF} \times s_{ig}R_t + \lambda_i + \lambda_t + e_{it}^{CF}. \tag{1.20}$$

Under the assumptions of this setting, this regression identifies β^P in one step as OLS partials-out the correlation between G_t and R_t. This estimation strategy is often referred to as the *control function approach*. In such a case, there is no need for the additional time-series step that I discuss next. However, the control function approach works under restrictive assumptions on the dynamics of the data generating process, as I clarify in the next subsection. Because these assumptions are hard to justify in macroeconomic environments, I outline the time series step to (i) show that the decomposition and control function approach yield equivalent results in static settings and (ii) ease the transition to dynamic settings.

Time-series Step

The purpose of this step is two-fold. First, it estimates the response of R_t to each aggregate shock. Second, it uses these responses to compute the size of a hypothetical ε_{rt} shock that matches the response of R to a unit change in ε_{gt}. The time series responses to each shock can be estimated using:

$$R_t = v\varepsilon_{gt} + a\varepsilon_{rt} + e_{rt}, \tag{1.21}$$

with the estimated coefficients converging to:

$$\mathbb{E}[\hat{v}] = \boldsymbol{R}_{g,t} = \delta, \qquad\qquad \mathbb{E}[\hat{a}] = \boldsymbol{R}_{r,t}. \tag{1.22}$$

Here $\hat{v}$ denotes the estimated response of R to a unit change in ε_{gt} and $\hat{a}$ the estimated response to a unit change in ε_{rt}. Next, I use these estimated responses to calculate a hypothetical shock ε_{rt} that changes R_t exactly by $\hat{v}$. This hypothetical shock— denoted by $\tilde{\varepsilon}_{rt}$ —is given by:

$$\hat{\tilde{\varepsilon}}_{rt} = \frac{\hat{v}}{\hat{a}} \quad \Rightarrow \quad \mathbb{E}[\hat{\tilde{\varepsilon}}_{rt}] = \frac{\delta}{\boldsymbol{R}_{r,t}}. \tag{1.23}$$

Final Step

The final step uses $\hat{c}$, the estimated reduced form response of Y_i to R from the cross-sectional step, to evaluate the effect of $\tilde{\varepsilon}_{rt}$, the hypothetical shock to R from the time series step. This yields a consistent estimate of Ω, the HEDGE term:

$$\mathbb{E}[\hat{\Omega}] = \mathbb{E}[\hat{c} \times \hat{\tilde{\varepsilon}}_{rt}] = \tilde{\gamma} \times \delta = \Omega. \tag{1.24}$$

A consistent estimate for the portable elasticity can be constructed as follows.

$$\hat{\beta}^P = \hat{b} - \hat{c} \times \hat{\tilde{\varepsilon}}_{rt} \quad \Rightarrow \quad \mathbb{E}[\hat{\beta}^P] = \beta^P, \tag{1.25}$$

where $\hat{b}$ and $\hat{c}$ are estimated in the cross-sectional step and $\hat{\tilde{\varepsilon}}_{rt}$ is estimated in the time series step. As mentioned above, under the assumptions made in this section, this three-step procedure yields an equivalent result to the control function approach.

Monte Carlo Simulations

The performance of the decomposition is illustrated using Monte Carlo simulations. I simulate data from an economy characterized by the system of equations presented above and estimate three different specifications: (i) Baseline, (ii) Decomposition, and (iii) Control function. Baseline corresponds to the TWFE regression. Decomposition corresponds to the results of applying the framework presented above and control function to the results from using specification (1.20). Figure 1.1 presents the distribution of the estimated coefficients for each strategy in an economy with $\beta = \gamma = \delta = \phi = .5$ and $\alpha = 0$. The population values for the portable and HEDGE terms are $\beta^P = .5$ and $\Omega = .125$, respectively. There are two takeaways. First, the baseline estimates are biased and centered on $\beta^P + \Omega = .625$. Second, both the decomposition and control function approach consistently estimate β^P. Figure A.1 in Appendix A.1.1 shows that the same results hold for an economy with two-way general equilibrium feedback between G and R (that is, $\alpha \neq 0$).

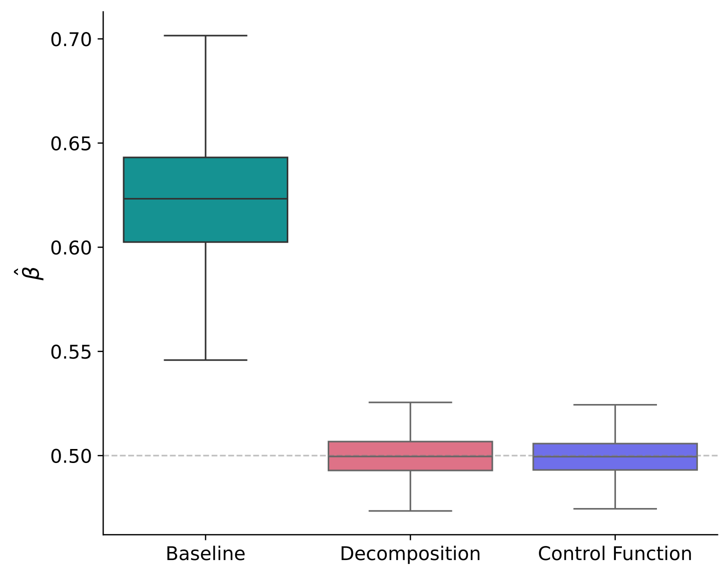


Figure 1.1. Static Monte Carlo Simulations - $\beta^P = .5$

Note: Distribution of point estimates for the *Baseline*, *Decomposition* and *Control Function* estimation strategies. Based on 1000 repetitions with sample size of $n = 100$ and $t = 300$ and parameters set to $\delta = \beta = \gamma = \phi = .5$.

1.3.2 Dynamic Markov Settings - No Anticipation Effects

This subsection studies dynamic environments where aggregate variables evolve according to a finite-order VAR, while cross-sectional outcomes respond to current and lagged realizations of these aggregates and may exhibit their own persistence. I refer to these as Markov environments.

These settings lie between static and forward-looking models, allowing for persistence while abstracting from anticipation effects.

In this environment, the control function approach generally fails to deliver consistent estimates, even though it is valid in static settings. The Markov structure makes it possible to isolate the source of this failure. It also provides a transparent benchmark that serves as a stepping stone to the forward-looking environments studied in Section 1.3.3.

In Markov environments, aggregate and cross-sectional variables evolve according to the following system of equations.

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \end{bmatrix},$$

$$Y_{it} = \sum_{s=0}^{S} \beta_s s_{ig} G_{t-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t-k} + \sum_{l=1}^{L} \psi_l Y_{i,t-l} + u_i + u_t + u_{it}^{y}, \tag{S2}$$

where G_t and R_t follow a $VAR(p)$ process, allowing for responses to each other's structural shocks u_{gt} and u_{rt}. The unit-level outcome Y_{it} follows an $AR(L)$ process and loads heterogeneously on contemporaneous and lagged realizations of both G_t and R_t, with the same exposure shifters applied uniformly across all lags. All other features of the environment are as in the static setting described in Section 1.3.1.

I assume that agents do not anticipate future realizations of aggregate shocks. Under this Markov structure, sequences of shocks that replicate a given path of aggregate variables ex post also reproduce the corresponding outcome responses.

For exposition purposes, I use as an example an economy with $P = 1, S = K = L = 0$ and where G does not respond to R. As in the static setting, I normalize the standard deviation of idiosyncratic shocks to unity. This yields familiar expressions that are helpful in building intuition. Formally, the economy is characterized by:

$$G_t = \rho_g G_{t-1} + u_{gt}, \tag{1.26}$$

$$R_t = \rho_r R_{t-1} + \delta G_t + u_{rt}, \tag{1.27}$$

$$Y_{it} = \beta \times s_{ig} G_t + \gamma \times s_{ir} R_t + u_i + u_t + u_{it}^y. \tag{1.28}$$

I refer to this as the $AR(1)$ case. This simplified example departs from the static setting only by introducing persistence in the aggregate variables. The parameters ρ_g and ρ_r govern the strength of persistence in G and R, respectively.

The goal is to estimate β_h^P, the dynamic differential response of Y_i to a shock to G_t, holding the time-path of R fixed. For the $AR(1)$ example, this implies:

$$\beta_h^P = \beta \times \rho_g^h \qquad \forall \ h, \tag{1.29}$$

where β_h^P captures the cumulative direct effect of a shock to G_t on Y_{it+h}, holding the entire path of R fixed. Consider a researcher who observes an innovation in G, denoted by ε_{gt}, that satisfies Assumption 1.3.

Assumption 1.3. *The observed innovation ε_{gt} is a noisy measure of the structural shock u_{gt}:*

$$\varepsilon_{gt} = \kappa u_{gt} + \eta_{gt}, \quad \kappa \neq 0, \tag{1.30}$$

where $\eta_{gt} \sim N(0, \sigma_v^2)$ is orthogonal to contemporaneous, past, and future structural innovations:

$$cov[\eta_{gt}, u_{ks}] = 0 \quad \forall \ k \in \{g, r\}, \ s. \tag{1.31}$$

Throughout, I denote the impulse response of X to an observed innovation ε_{kt} by $\boldsymbol{X}_k = \{\boldsymbol{X}_{k,h}\}_h$. Here $\boldsymbol{X}_{k,h}$ denotes the impulse response of X to ε_{kt} at horizon h. As before, I normalize κ such that $\boldsymbol{G}_{g,0} = 1$.

The TWFE estimation strategy can be implemented through the following panel local projection for each horizon h:

$$Y_{it+h} = b_h \times s_{ig} G_t + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{1.32}$$

where $s_{ig} G_t$ is instrumented with $s_{ig}\varepsilon_{gt}$ to address potential endogeneity concerns in G_t. Here, b_h measures the cumulative effect of a change in G_t on Y_{it+h} for the unit with average exposure. Horizon-specific unit and time fixed effects are captured by λ_{ih} and λ_{th}, respectively. The estimate for b_h converges to:

$$\mathbb{E}[\hat{b}_h] = \underbrace{\beta_h^P}_{\text{Portable Effect}} + \underbrace{\Omega_h}_{\text{HEDGE term}}. \tag{1.33}$$

The first term is the portable component that measures the effect of interest, holding the path of R fixed. The second term Ω_h captures the HEDGE component that is now horizon-specific. In the AR(1) case, these terms can be expressed as:[15]

$$\begin{aligned} \mathbb{E}[\hat{b}_h] &= \beta\, \boldsymbol{G}_{g,h} + \gamma\phi \boldsymbol{R}_{g,h} \\ &= \underbrace{\beta \times \rho_g^h}_{\beta_h^P} + \underbrace{\gamma\phi \times \delta \sum_{j=0}^{h} \rho_r^{h-j} \rho_g^j}_{\Omega_h}. \end{aligned} \tag{1.34}$$

If $\phi = 0$ or $\delta = 0$, the HEDGE terms vanish and the benchmark strategy recovers the true portable effect. We can break down Ω_h into two terms:

$$\Omega_h = \tilde{\gamma} \times \boldsymbol{R}_{g,h}. \tag{1.35}$$

[15]The assumptions on ε_{gt} imply that the first-stage coefficient $\boldsymbol{G}_{g,0} = 1$, so I omit it for notational simplicity.

As in the static setting, the first term $\tilde{\gamma} = \gamma \times \phi$ measures the differential response of Y_i to a change in R, conditional on s_{ig}. The second term measures the dynamic response of R to a shock to G.

Next, I outline how to implement each step of the decomposition to estimate Ω_h in Markov settings. The additional input relative to the TWFE strategy is the same as in the static setting: an instrument for the contemporaneous realization of R. In what follows, I assume that, on top of the shock of interest ε_{gt}, we observe a contemporaneous innovation to R_t, denoted by ε_{rt}, that satisfies Assumption 1.4.

Assumption 1.4. *The observed innovation ε_{rt} is a noisy measure of the structural shock u_{rt}:*

$$\varepsilon_{rt} = \upsilon u_{rt} + \eta_{rt}, \quad \upsilon \neq 0, \tag{1.36}$$

where $\eta_{rt} \sim N(0, \sigma^2_{\eta_r})$ is orthogonal to contemporaneous, past, and future structural innovations:

$$cov[\eta_{rt}, u_{ks}] = 0 \quad \forall\, k \in \{g, r\},\ s. \tag{1.37}$$

This assumption implies that ε_{rt} isolates exogenous variation in R_t that is orthogonal to contemporaneous, past, and future innovations to G, as well as, past and future innovations to R.[16]

Cross-sectional Step

The goal of this step is to identify the dynamic reduced-form responses of Y_i to G and R, conditional on s_{ig}. For identification of the R responses, we exploit the variation in R induced by the innovation ε_{rt}. The following local projection can be used to estimate these reduced-form

[16] This assumption allows for correlation between ε_{gt} and ε_{rt}. The HEDGE term can still be identified in this case as OLS partials out any shared variation.

responses:

$$Y_{it+h} = b_h \times s_{ig}\varepsilon_{gt} + c_h \times s_{ig}\varepsilon_{rt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{1.38}$$

where each aggregate shock is interacted with s_{ig}, the exposure shifter to G. First, this specification identifies the same b_h as the baseline regression.[17] This is because the additional interaction only captures variation in R that is not correlated with G. For the AR(1) case, this implies

$$\mathbb{E}[\hat{b}_h] = \beta\, \boldsymbol{G}_{g,h} \;+\; \tilde{\gamma}\; \boldsymbol{R}_{r,h}. \tag{1.39}$$

Second, this regression also identifies the response of Y_i to R, conditional on s_{ig} *and* conditional on the path of R after a change in ε_{rt}:

$$\mathbb{E}[\hat{c}_h] = \tilde{\gamma}\; \boldsymbol{R}_{r,h}. \tag{1.40}$$

Inspection of $\hat{b}_h$ and $\hat{c}_h$ clarifies that the omitted variable bias due to HEDGE is a function of $\boldsymbol{R}_g$, while the estimated reduced-form response of Y_i to R is a function of a different source of variation, namely $\boldsymbol{R}_r$.

In this example, expressing the impulse responses in terms of parameters yields:

$$\mathbb{E}[\hat{b}_h] = \underbrace{\beta \times \rho_g^h}_{\beta_h^P} + \underbrace{\tilde{\gamma} \times \delta \sum_{j=0}^{h} \rho_r^{h-j} \rho_g^j}_{\Omega_h}, \qquad \mathbb{E}[\hat{c}_h] = \tilde{\gamma} \times \rho_r^h. \tag{1.41}$$

Here, the estimated coefficient on the interaction term $s_{ig}\varepsilon_{rt}$ is a function of the differential sensitivity parameter $\tilde{\gamma}$ and the auto-regressive structure of R. To better understand the link

[17] In the general case in which the first-stage coefficient on ε_{gt} is different from one, then the reduced form regression identifies a rescaled version of the expression in (1.34). The rescaled version is given by $\kappa_g \times \hat{b}_h^{TW}$ where κ_g is the coefficient of the first stage and $\hat{b}_h^{TW}$ is the *2SLS* coefficient of (1.34). Given that the decomposition results are robust to arbitrary first-stage values, I omit them from the main outline. The important point is that introducing the interaction between s_{ig} and ε_{rt} does not directly address the omitted variable bias due to HEDGE.

between $\hat{c}_h$ and Ω_h, we can rewrite the latter as:

$$\Omega_h = \left[\tilde{\gamma}\rho_r^h + \tilde{\gamma}\sum_{j=1}^{h}\rho_r^{h-j}\rho_g^j\right]\times\delta \qquad \forall\ h \geq 0. \tag{1.42}$$

Consider the special case where G_t is an impulse (i.e. $\rho_g = 0$) so that the second term in the expression above drops:

$$\Omega_h = \tilde{\gamma}\rho_r^h \times \delta. \tag{1.43}$$

In this case, $\hat{c}_h$ identifies a rescaled version of Ω_h, where the scale depends on δ, the parameter that governs the contemporaneous response of R_t to G_t:

$$\Omega_h = \tilde{\gamma}\rho_r^h \times \delta \qquad vs. \qquad \mathbb{E}[\hat{c}_h] = \tilde{\gamma}\rho_r^h. \tag{1.44}$$

Intuitively, this scaling captures that in this example a change of one unit in ε_{rt} changes R_t by one unit, whereas this change is size δ for a change of one unit in ε_{gt}. Therefore, the only piece of information missing to compute Ω_h in this scenario is δ. Or, more generally, the ratio between the impact response of R to each aggregate shock, as in the static setting example.

Next, the time-series step provides a general strategy to take into account the difference between the impulse responses $\boldsymbol{R}_g$ and $\boldsymbol{R}_r$, which serves as a bridge between the HEDGE term and $\hat{c}_h$.

Time-series Step

This step provides the additional inputs required to move from $\hat{c}_h$ to $\hat{\Omega}_h$. From the previous step, we know how to evaluate the cross-sectional propagation of changes in R caused by ε_{rt} using $\hat{c}_h$. Following the methodology of Sims and Zha (2006), one can find a combination of innovations in R, denoted by $\tilde{\boldsymbol{\varepsilon}}_r = \{\tilde{\varepsilon}_{rt+h}\}_h$, which taken together reproduce $\boldsymbol{R}_g$ the impulse response of R to a G shock. Then we can use the reduced-form estimates from the cross-sectional step to evaluate the response of Y_i to the sequence of innovations $\tilde{\boldsymbol{\varepsilon}}_r$.

By construction, this sequence of innovations exactly replicates $\boldsymbol{R}_g$ in the *ex post* sense. This is because at each point in time $h > 0$ we hit the economy with additional surprises to R, *unexpected* from the perspective of $h = 0$. Under the Markov assumption, innovations are not anticipated, so enforcing the path $\boldsymbol{R}_g$ ex post through a sequence of unexpected shocks yields the same outcome response as if that path had been anticipated ex ante. Section 1.3.3 relaxes the assumption of no anticipation effects following the work by McKay and Wolf (2023) and Barnichon and Mesters (2023).

Applying the approach of Sims and Zha (2006) requires estimates of the impulse response of R to the two aggregate shocks, ε_{gt} and ε_{rt}. These can be estimated using traditional time series methods, for example, local projections (Jordà, 2005):

$$R_{t+h} = v_h \varepsilon_{gt} + a_h \varepsilon_{rt} + e_{rt+h}. \tag{1.45}$$

The regression coefficients $\hat{v}_h$ and $\hat{a}_h$ converge to the following expression:

$$\mathbb{E}[\hat{v}_h] = \boldsymbol{R}_{g,h}, \qquad\qquad \mathbb{E}[\hat{a}_h] = \boldsymbol{R}_{r,h}. \tag{1.46}$$

In the AR(1) example, this is equivalent to:

$$\mathbb{E}[\hat{v}_h] = \delta \rho_r^h + \delta \sum_{j=1}^{h} \rho_r^{h-j} \rho_g^j, \qquad\qquad \mathbb{E}[\hat{a}_h] = \rho_r^h. \tag{1.47}$$

In this example, unless $\rho_g = 0$, the path of R will differ between shocks even after rescaling by δ.

We can use $\hat{\boldsymbol{R}}_g = \{\hat{v}_h\}_h$ and $\hat{\boldsymbol{R}}_r = \{\hat{a}_h\}_h$ to find the sequence of innovations $\tilde{\boldsymbol{\varepsilon}}_r$. Before presenting the general formula, I illustrate how the procedure works using the AR(1) example.

Finding the innovation for period $h = 0$ requires:

$$\hat{a}_0 \hat{\tilde{\varepsilon}}_{rt+0} = \hat{v}_0 \qquad \Rightarrow \qquad \mathbb{E}[\hat{\tilde{\varepsilon}}_{rt+0}] = \mathbb{E}\left[\frac{\hat{v}_0}{\hat{a}_0}\right] = \delta. \tag{1.48}$$

where $\hat{\tilde{\varepsilon}}_{rs}$ denotes the estimated counterfactual shock for period s. In other words, we look for the shock size to R that matches the estimated impact effect of a unit change in ε_{gt}. From the perspective of $h = 1$, this innovation implies that R changes by:

$$\mathbb{E}[\hat{a}_1 \times \hat{\tilde{\varepsilon}}_{rt+0}] = \delta\rho_r \qquad \neq \quad \mathbb{E}[\hat{v}_1] = \delta\rho_r + \delta\rho_g. \tag{1.49}$$

Although the innovation for period $h = 0$ matches the impact response $\hat{v}_0$, it does not match the response at $h = 1$ that is given by $\hat{v}_1$. Therefore, we feed in an additional innovation that hits the economy in period $h = 1$ to account for the remaining difference between the two paths:

$$\hat{v}_1 = \hat{a}_1 \times \hat{\tilde{\varepsilon}}_{rt+0} + \hat{a}_0 \times \hat{\tilde{\varepsilon}}_{rt+1}, \tag{1.50}$$

where to evaluate the effect of this second innovation we use the contemporaneous coefficient $\hat{a}_0$. This implies that the estimated $\tilde{\varepsilon}_{rt+1}$ satisfies:

$$\mathbb{E}[\hat{\tilde{\varepsilon}}_{rt+1}] = \delta\rho_g. \tag{1.51}$$

The sequence of ex-post innovations that replicates the path $\hat{\boldsymbol{R}}_g$ up to horizon H is characterized by Proposition 1.1.

Proposition 1.1. *Under Assumption 1.4, $a_0 \neq 0$. Hence, for any horizon H, there exists a unique* *sequence of innovations* $\{\tilde{\varepsilon}_{rt+h}\}_{h=0}^{H}$ *such that*

$$v_h = \sum_{j=0}^{h} a_{h-j} \tilde{\varepsilon}_{rt+j} \quad \forall\, h \in \{0, \ldots, H\}. \tag{1.52}$$

Moreover, since $\hat{a}_0 \xrightarrow{p} a_0$*, it follows that* $\hat{a}_0 \neq 0$ *with probability approaching one, so that the empirical analogue*

$$\hat{v}_h = \sum_{j=0}^{h} \hat{a}_{h-j} \hat{\tilde{\varepsilon}}_{rt+j} \tag{1.53}$$

is well-defined.

This result applies to any economy that satisfies (S2), including settings with two-way feedback between G and R.

Final Step

Once we estimate $\tilde{\boldsymbol{\varepsilon}}_r$, we can compute $\hat{\Omega}_h$ as follows:

$$\hat{\Omega}_h = \sum_{k=0}^{h} \hat{c}_{h-k} \hat{\tilde{\varepsilon}}_{rt+k} \qquad \forall\, h \in H. \tag{1.54}$$

That is, we estimate $\hat{\Omega}_h$ using the reduced-form estimates from the cross-sectional step to evaluate the propagation of the sequence of innovations $\tilde{\boldsymbol{\varepsilon}}_r$. Crucially, the cross-sectional estimates $\hat{c}_h$ are evaluated using a sequence of shocks that generate the same variation in R that we used to identify those cross-sectional coefficients in the first place. Under the assumptions of this subsection, this yields a consistent estimate of the HEDGE term at each horizon h. Lastly, the portable elasticity estimate is computed as follows.

$$\hat{\beta}_h^P = \hat{b}_h - \hat{\Omega}_h. \tag{1.55}$$

where $\hat{b}_h$ is the estimated coefficient on $s_{ig}\varepsilon_{gt}$ from the cross-sectional step.

Control Function Approach in Dynamic Settings

Inclusion of the interaction $s_{ig}R_t$ as a control, in general, does not deliver a consistent estimate of β_h^P. The reason is that the additional interaction term can only control for time paths of R that can be spanned by variation in the included controls, instead of variation in ε_{gt}. Because these two sources of variation may generate different time paths for R, the control function approach is not a valid robustness check for HEDGE in general macroeconomic environments. Intuitively, the dynamic response of R depends on the shock that drives it. For example, monetary policy may respond to a demand shock with a hump-shaped path due to inertia, while orthogonal fluctuations in R may follow a more AR(1)-like process. A contemporaneous control $s_{ig}R_t$ cannot distinguish between these state-contingent paths.

Concretely, the control function approach specification is given by:

$$Y_{it+h} = b_h^{CF} \times s_{ig}G_t + c_h^{CF} \times s_{ig}R_t + \lambda_{ih}^{CF} + \lambda_{th}^{CF} + e_{it+h}^{CF}, \tag{1.56}$$

where $s_{ig}G_t$ is instrumented with $s_{ig}\varepsilon_{gt}$, as in the baseline TWFE strategy, and the superscript CF is used to denote the estimates from the control function approach. In the AR(1) example, it can be shown that:

$$\mathbb{E}[\hat{b}_h^{CF}] - \beta_h^P \propto \tilde{\gamma} \times \Big(\boldsymbol{R}_{g,h} - \delta \boldsymbol{R}_{\tilde{R},h}\Big), \tag{1.57}$$

where $\delta\, \boldsymbol{R}_{\tilde{R},h}$ is proportional to the path of R spanned by the included controls (in this case, the contemporaneous realization R_t).[18] Specifically, δ, captures the contemporaneous

[18]Plugging in the expressions for $\boldsymbol{R}_{g,h}$ and $\boldsymbol{R}_{\tilde{R},h}$ yields:

$$\boldsymbol{R}_{g,h} - \delta \boldsymbol{R}_{\tilde{R},h} = \delta \rho_r^h + \delta \sum_{j=1}^{h} \rho_r^{h-j} \rho_g^j - \delta \underbrace{\rho_r^h}_{\boldsymbol{R}_{\tilde{R},h}} = \delta \sum_{j=1}^{h} \rho_r^{h-j} \rho_g^j. \tag{1.58}$$

In this AR(1) example, the control function approach partially addresses the problem by taking care of the correlation between ε_{gt} and R_{t+h} resulting from the pure auto-regressive structure of R. However, it cannot account for the dynamic effects of ε_{gt} on R_{t+h} that go through $\{G_{t+k}\}_{k=1}^{h}$. Note that, for example, if R responded to G with a lag then the control function approach would, in expectation, yield no improvements over the TWFE strategy for $h > 0$.

response of R to the G shock.

Expression (1.57) highlights that, unless $\boldsymbol{R}_{g,h} = \delta \boldsymbol{R}_{\tilde{R},h}$, the control function approach does not provide a consistent estimate of β_h^P. For example, if the path of R resembles an AR(2) conditional on ε_{gt} but an AR(1) unconditionally then $\boldsymbol{R}_{g,h} \neq \delta \boldsymbol{R}_{\tilde{R},h}$. In Appendix A.1.3, I derive the expected value of $\hat{b}_h^{CF}$ for alternative DGPs to further illustrate the mechanics of the control function approach. Moreover, while the control function approach is inconsistent in general, its performance relative to its (biased) population value can vary substantially with model mis-specification. In particular, when the set of additional controls, such as lagged dependent variables or lagged shocks, is misspecified, the resulting estimates can deviate considerably from that population value.

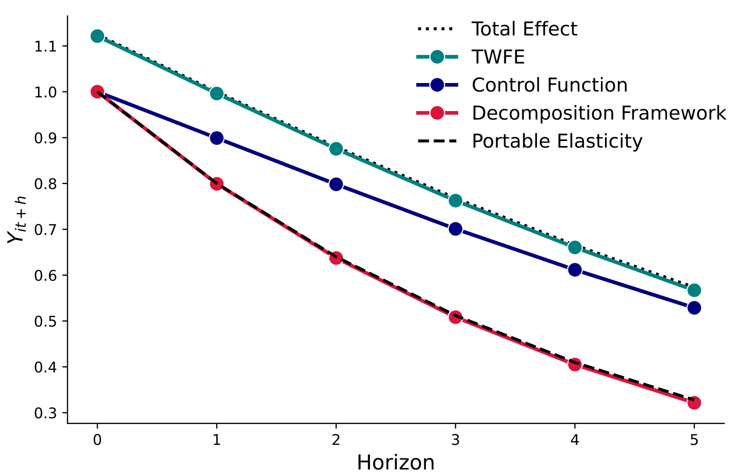


Figure 1.2. Dynamic Setting Without Anticipation Effects - Monte Carlo Results.

Note: Solid lines show the estimated cross-sectional responses to a G shock for three specifications - *TWFE, Decomposition Framework* and *Control Function* - together with the population portable and total (β_h^P + Ω_h) elasticities. Shaded areas show 90% confidence bands. Based on 1000 repetitions with $N = 300$ and $T = 1000$. The DGP is generated using the same parameters as for the static case plus $\rho_g = .8, \rho_r = .8, \rho_y = .0$.

Monte Carlo Simulations

The performance of the decomposition is illustrated through a series of Monte Carlo simulations. I simulate data for different economies that satisfy the Markov assumptions of system (S2) and estimate three different impulse response functions: i) TWFE, ii) Control

Function, iii) Decomposition Framework. These correspond respectively to specifications (1.32) and (1.56), and to the portable elasticity obtained from the decomposition strategy discussed above.

I present two sets of results. First, in Figure 1.2 I show results for the AR(1) example. Full details on the parameters used for the simulation can be found in the Appendix A.1.3. The figure shows the estimated cross-sectional responses to ε_{gt}. The dashed black line corresponds to the population portable elasticity, while the dotted black line corresponds to the population total effect (i.e., $\beta_h^P + \Omega_h$). The decomposition framework provides a consistent estimate of the portable elasticity at all horizons. In contrast, neither the TWFE strategy nor the control function approach identify the portable elasticity.

Second, I study the performance of the decomposition and control function approaches in a variety of data-generating processes that fit into (S2). The sample size is set to $N = 300$ and $T = 1000$, and the number of repetitions to $J = 1000$. Appendix A.1.3 details the set of data-generating processes and their parameter values. For each estimation strategy k, DGP z and horizon h, I compute the absolute difference between the population portable elasticity, β_{zh}^P, and the average estimate across all repetitions and then normalize it by β_{zh}^P to make it comparable across DGPs.

$$\theta_{zh}^k = \frac{\mid J^{-1} \sum_{j=1}^{J} \hat{b}_{jzh}^k - \beta_{zh}^P \mid}{\mid \beta_{zh}^P \mid}. \tag{1.59}$$

Consistency requires $\theta_{zh}^k \to 0$. Table 1.1 shows the mean and standard deviation of θ_{zh}^k by horizon and estimation method. The decomposition approach consistently recovers the true portable effects as opposed to the control function approach and TWFE baseline strategy.

Extensions

Appendix A.1.3 presents the derivations and results for settings with two-way general equilibrium feedback (2W-GE) between aggregate variables. This analysis yields an important

Table 1.1. Absolute Mean Relative Bias by Estimation Method and Horizon.

Estimation Strategy	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
Decomposition	0.002	0.002	0.007	0.013	0.021	0.031
	(0.003)	(0.002)	(0.007)	(0.014)	(0.023)	(0.035)
Control Function	0.445	0.259	0.245	0.282	0.379	0.516
	(0.494)	(0.223)	(0.182)	(0.300)	(0.417)	(0.519)
TWFE	0.146	0.226	0.333	0.463	0.603	0.743
	(0.130)	(0.192)	(0.295)	(0.385)	(0.480)	(0.580)

Note: Standard errors shown in parentheses. Values are reported as fractions of the population portable elasticity (1 = 100 %). The sample size is set to $N = 300$ and $T = 1000$ and the number of repetitions to $J = 1000$. See Appendix A.1.3 for further details on each parametrization.

insight. I show that, in general, when there is 2W-GE, the cross-sectional elasticities estimated using either two-stage least squares (2SLS) or reduced-form analysis will be contaminated by GE effects, even in the absence of HEDGE. In other words, in such settings, the HEDGE conditions discussed at the beginning of the paper are sufficient to break identification but not necessary. The TWFE strategy will, in general, not identify a clean partial equilibrium elasticity in these settings. This happens because units that are more exposed to G will be more exposed to both the *autonomous* change in G triggered by ε_{gt} and to the posterior change in response to R. This is an overlooked source of omitted variable bias that this paper brings to light and which the decomposition framework *also* addresses.

Appendix A.1.3 shows that the decomposition framework can accommodate settings where both R_t and ε_{gt} are correlated with a third aggregate variable A_t (for example, TFP, exchange rate), but A_t affects all units in the same way. This echoes empirical applications for which the available source of variation ε_{gt} is not as good as random at the aggregate level, but it does satisfy exogeneity with respect to A_t in the cross section. Note that this is equivalent to saying that there is no HEDGE with respect to A_t.

Appendix A.1.4 discusses three alternative estimation strategies that one might expect to address HEDGE, and demonstrates through a Monte Carlo simulation exercise that they fail to do so. The strategies considered are (i) interactive fixed effects (Bai, 2009); (ii) alternative version of the CFA including $s_{ir}R_t$ rather than $s_{ig}R_t$; (iii) alternative CFA instrumenting $s_{ig}R_t$ with $s_{ig}\varepsilon_{rt}$.

1.3.3 Forward-Looking Dynamic Setting

This section generalizes the decomposition framework to settings with forward-looking dynamics. This is akin to the dynamics implied by DSGE models, which I use as a laboratory to test the performance of the framework. The difference with respect to the preceding section is that now there can be anticipation effects: agents respond both to current and expected realizations of macroeconomic variables. The key implication is that estimating the differential response to the conditional path of the GE variable, R, requires feeding the system a counterfactual sequence of R shocks that reproduces $\boldsymbol{R}_g$ both ex post and in expectation.

As shown by McKay and Wolf (2023) and Barnichon and Mesters (2023), constructing counterfactuals that are robust to forward-looking dynamics requires access to shocks that shift both the contemporaneous and expected future realization of macroeconomic variables. First, I show how to implement the decomposition when a researcher has access to these types of shocks and then I discuss implementation in cases where access to shocks is limited. In settings with limited access to news shocks, the decomposition framework provides an approximation to the population portable elasticity. The conditions that determine the precision loss resulting from limited information depend on the data-generating process and the shocks under analysis. As an illustration, in Section 1.5.3, I study the determinants of the decomposition accuracy in the context of a two-region New Keynesian model used to estimate cross-sectional fiscal multipliers.

Consider an economy in which the unit-level outcome can be characterized as follows.

$$\begin{aligned} Y_{it} = \sum_{f=0}^{F} \beta_f s_{ig} \mathbb{E}_t[G_{t+f}] + \sum_{f=0}^{F} \gamma_f s_{ir} \mathbb{E}_t[R_{t+f}] \\ + \sum_{s=1}^{S} \beta_s^{\ell} s_{ig} G_{t-s} + \sum_{k=1}^{K} \gamma_k^{\ell} s_{ir} R_{t-k} + \sum_{l=1}^{L} \psi_l Y_{it-l} + u_i + u_t + u_{it}^{y}, \end{aligned} \tag{1.60}$$

where $\mathbb{E}_t[X_{t+f}]$ is the expectation at time t for the realization of X in period $t+f$. The parameter

β_f, for $f > 0$, captures the present response to the expected realization of G in period $t+f$. Similarly for γ_f. I assume that only the first F expected realizations of aggregate variables matter for the present responses and use the superscript ℓ to denote the responses to lagged realizations of the aggregate variables.[19] In addition, I assume that exposure shifters are variable-specific but not horizon specific (i.e., s_{ig} determines the cross-sectional response to both contemporaneous, expected, and lagged realizations of G).[20]. The aggregate processes are as in the Markov economy with the addition that they are subject to both contemporaneous and news (or anticipated) structural innovations:

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} \boldsymbol{u}_{gt} \\ \boldsymbol{u}_{rt} \end{bmatrix}, \tag{1.61}$$

where $\boldsymbol{u}_{kt}$ is the vector of contemporaneous and news shocks to k in period t.[21]

Assumption 1.5. *For each aggregate variable $k \in \{g, r\}$, the innovation vector $\boldsymbol{u}_{kt}$ admits the decomposition*

$$\boldsymbol{u}_{kt} = \{u^t_{kt+f}\}_{f=0}^{F}, \tag{1.62}$$

where u^t_{kt} is a contemporaneous shock and, for $f \geq 1$, u^t_{kt+f} is a news shock revealed at time t about the realization of variable k at time $t+f$.

[19] Alternatively, the DGP in (1.60) can be interpreted as an approximation of a true DGP with $F^* = \infty$ and $\gamma_f \approx 0$ for $f > F$.

[20] The estimation framework does not require knowledge of the true exposure shifters to R so assuming a constant s_{ir} is without loss of generality. Horizon-specific shifters to R, $s_{ir,j}$ for $j \in \{F, K\}$, imply that ϕ becomes horizon-specific which will be reflected in the estimated cross-sectional responses to different news shocks (See cross-sectional step below). The assumption that s_{ig} is a good proxy for exposure to present and expected realizations of G is in line with most empirical work. If the exposure shifters to the contemporaneous realization of G, s_{ig}, and to the expected realizations of G, $s_{ig,f}$, differ, then $s_{ig} \times \varepsilon_{gt}$ will be a strong instrument for $s_{ig} \times G_t$ but possibly weaker for $s_{ig} \times \mathbb{E}_t[G_{t+f}]$. This is independent of HEDGE and, therefore, relaxing the assumption of constant s_{ig} exceeds the scope of this paper.

[21] For example, the first entry of $\boldsymbol{u}_{kt}$ corresponds to a contemporaneous shock, while the second one corresponds to news about the realization of k in period $t+1$.

All innovations are mean zero and mutually orthogonal:

$$cov[u^t_{kt+f}, u^t_{k't+f'}] = 0 \quad \textit{for all } (k, f, t) \neq (k', f', t'). \tag{1.63}$$

This orthogonality assumption implies that each news shock shifts expectations about a specific future realization of the aggregate variable without affecting other horizons or contemporaneous innovations. The rest of the economy is the same as in the Markov setting of Subsection 1.3.2.

The object of interest is the dynamic effect of a shock to G_t on the cross-sectional outcome, holding the realized and *expected* path of R fixed. Consider the following panel local projection with two-way fixed effects:

$$Y_{it+h} = b_h \times s_{ig} G_t + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{1.64}$$

where $s_{ig}\varepsilon_{gt}$ is used as an instrument for $s_{ig}G_t$ to address potential endogeneity concerns. To simplify notation, I assume that $\beta_j = \beta \; \forall \; f, \ell$ and shut down the parameters on lagged realizations of aggregate and local variables in the process for Y_{it}. Then, $\hat{b}_h$ converges to the following expression:

$$\mathbb{E}[\hat{b}_h] = \beta^P_h + \underbrace{\sum_{f=0}^{F} \gamma_f \phi \frac{cov[\mathbb{E}_t[R_{t+h+f}], \varepsilon_{gt}]}{v[\varepsilon_{gt}]}}_{\Omega_h}, \tag{1.65}$$

where the first term is the portable component that captures the effect of a shock to G_t on Y_{it+h}, including the effects through changes in future G, holding $\mathbb{E}_t[R_{t+h}]$ for $h \geq 0$ constant. The second term is the HEDGE component which is a function of (i) the contemporaneous and expected changes in R_t conditional on a shock to G_t and (ii) the cross-sectional sensitivities to contemporaneous and expected future realizations of R, conditional on s_{ig}. Next, I outline how the cross-sectional and time-series steps can be generalized to accommodate forward-looking dynamics.

Cross-sectional Step

This section outlines the estimation strategy used to identify the dynamic reduced-form responses of Y_i to G_t and to both current and expected changes in R. The latter set of parameters measures the response of Y_{it+h} to the expectation formed at time t about the realization of R_{t+f}. For example, for $h = f = 0$, this parameter measures the contemporaneous response of Y_{it} to a contemporaneous innovation in R. Identifying these parameters requires exogenous variation in the period-t expectation about the realization of R_{t+f}. In other words, we need to observe news (or anticipated) shocks about the realization of R_{t+f} that agents learn about in period t. Let $\varepsilon^t_{r,t+f}$ denote an observed innovation that captures news received at time t about the realization of R_{t+f}. Analogously to the Markov setting, I assume that these observed innovations are noisy measures of the corresponding structural news shocks:

$$\varepsilon^t_{r,t+f} = \upsilon_f u^t_{rt+f} + \eta^t_{r,t+f}, \quad \upsilon_f \neq 0, \tag{1.66}$$

where the noise terms $\eta^t_{r,t+f}$ are orthogonal to all structural innovations. For example, an unexpected monetary policy announcement this quarter about a possible one-period interest rate hike next quarter could be used to identify the coefficients for $f = 1$, at different horizons.

Suppose that we have access to news shocks up to horizon F, then we can estimate the following reduced-form local projections:

$$Y_{it+h} = b_h \times s_{ig}\varepsilon_{gt} + \sum_{f=0}^{F} c_{fh} \times s_{ig}\varepsilon^t_{r,t+f} + \lambda_{ih} + \lambda_{th} + e_{it+h}. \tag{1.67}$$

Here $\hat{c}_{f,h}$ is the reduced form response of Y_{it+h} to the news received at t about the realization of R_{t+f}, conditional on s_{ig}. As before, $\hat{b}_h$ is the reduced-form response of Y_{it+h} to a change in G_t, conditional on s_{ig}.

Time-Series Step

The time-series step estimates the impulse response of R to (i) the shock of interest, and (ii) the same set of contemporaneous and news shocks to R used in the cross-sectional step.

$$R_{t+h} = v_h \varepsilon_{gt} + \sum_{f=0}^{F} a_{fh} \varepsilon^t_{r,t+f} + e_{rt+h}, \tag{1.68}$$

where $\varepsilon^t_{r,t+f}$ is the news shock to R defined above. The coefficients a_{fh} measure the response of R_{t+h} to the information received in period t about the realization of R_{t+f}. Under rational expectations, $\hat{v}_h$ and $\hat{a}_{fh}$ are unbiased estimates of $cov[\mathbb{E}_t[R_{t+h}], \varepsilon_{gt}]$ and $cov[\mathbb{E}_t[R_{t+h}], \varepsilon^t_{r,t+f}]$, respectively.

Next, I use the impulse responses to the $F+1$ news shocks to find the combination of date-t shocks that replicates the estimated impulse response of R to ε_{gt}, $\hat{\boldsymbol{R}}_g = \{\hat{v}_h\}_h$. Let $\hat{A}$ be an $(H+1) \times (F+1)$ matrix whose (h, f)-th entry is $\hat{a}_{hf}$, the estimated response of R_{t+h} to a time-t news shock about R_{t+f}.

The estimated sequence of counterfactual date-t news shocks $\hat{\tilde{\boldsymbol{\varepsilon}}}^t = \{\hat{\tilde{\varepsilon}}^t_{r,t+f}\}_{f=0}^{F}$ is defined as the solution to:

$$\hat{\boldsymbol{R}}_g = \hat{A} \hat{\tilde{\boldsymbol{\varepsilon}}}^t. \tag{1.69}$$

Proposition 1.2.

i. If $\hat{\boldsymbol{R}}_g$ lies in the column space of $\hat{A}$, then there exists a sequence $\hat{\tilde{\boldsymbol{\varepsilon}}}^t$ that exactly replicates $\hat{\boldsymbol{R}}_g$. Moreover, if $\hat{A}$ has full column rank, this solution is unique.

ii. If $\hat{\boldsymbol{R}}_g$ does not lie in the column space of $\hat{A}$, the solution is given by the least squares projection:

$$\hat{\tilde{\boldsymbol{\varepsilon}}}^t = \arg \min_{e \in \mathbb{R}^{F+1}} \left\| \hat{\boldsymbol{R}}_g - \hat{A} e \right\|^2. \tag{1.70}$$

iii. *A sufficient condition for the uniqueness of the solution in (ii) is that $\hat{A}$ has full column rank, in which case:*

$$\hat{\tilde{\boldsymbol{\varepsilon}}}^t = (\hat{A}'\hat{A})^{-1}\hat{A}'\hat{\boldsymbol{R}}_g. \tag{1.71}$$

The rank condition on $\hat{A}$ is an empirical property of the estimated impulse responses. In particular, while Assumption 1.5 provides a structural interpretation of the shocks, uniqueness of the counterfactual decomposition depends on whether the observed news shocks span sufficiently rich variation in the data. The system is overidentified when $H > F$, exactly identified when $H = F$, and underidentified when $H < F$. In typical applications, $H > F + 1$, so an exact replication is not guaranteed. In this case, the solution corresponds to the projection of $\hat{\boldsymbol{R}}_g$ onto the space spanned by the news shocks observed at time t.

Final Step

Lastly, I compute $\hat{\Omega}_h$ by using the estimates from the cross-sectional step to evaluate the propagation of the estimated sequence of news shocks $\tilde{\boldsymbol{e}}^t$:

$$\hat{\Omega}^t_h = \sum_{f=0}^{F} \hat{c}_{fh} \times \hat{\tilde{\varepsilon}}^t_{r,t+f}, \tag{1.72}$$

where the superscript t is used to denote the estimated HEDGE term of the ex-ante decomposition. The estimate for β^P_h follows from the following:

$$\hat{\beta}^{P,t}_h = \hat{b}_h - \hat{\Omega}^t_h. \tag{1.73}$$

Limited News Shocks

In the cases where we observe $\tilde{F} < F$ news shocks, the decomposition yields an approximation to the population portable elasticities. In general, the loss of precision increases in the strength of forward-looking dynamics beyond $\tilde{F}$ (i.e., the importance of the parameters γ_f for

$f > \tilde{F}$) and in the difference between the observed path of R after the shock of interest and the path enforced by the estimated sequence of shocks $\tilde{\boldsymbol{e}}^t$. In Section 1.5.3, I use a two-region New Keynesian model to explore which observable moments can be used to study the performance of the framework in settings with limited access to news shocks.

DSGE Monte Carlo Simulations

To assess the performance of the decomposition framework, I conduct DSGE Monte Carlo (Ramey, 2016) simulations using data generated from a two-region New Keynesian model with government spending shocks. The model extends Nakamura and Steinsson (2014) by allowing heterogeneity across regions in both their exposure to government spending, G_t, and their sensitivity to the interest rate, R_t. Without loss of generality, I assume that $cov[s_{ig}, s_{ir}] < 0$, so that the region more exposed to G_t is less sensitive to R_t. Government spending follows an AR(1) process with innovation ε_{gt}, while the nominal interest rate, i_t, is determined by a Taylor rule and subject to its own shock ε_{it}. As a result, the real interest rate R_t responds positively and persistently to fiscal shocks (that is, $cov[R_{t+h}, \varepsilon_{gt}] > 0$). This calibration implies a positive HEDGE term. Appendix A.4 details the full model, and Table A.9 summarizes the calibration.

The simulated data are used to estimate cross-sectional fiscal multipliers using the estimation strategies discussed above. Given the forward-looking nature of standard New-Keynesian dynamics, I assume access to a finite number of news shocks (i.e., $\tilde{F} < F = \infty$) when applying the ex-ante decomposition.

I compare two parametrizations that differ in how R_t responds to a monetary shock ε_{it}, while holding fixed the response of R_t to ε_{gt}. In both scenarios, fiscal shocks induce a persistent increase in R_t. In the first scenario—labeled *misaligned*—monetary shocks induce only a transitory response in R_t, while in the second—labeled *aligned*—the response is calibrated to more closely mirror the one following fiscal shocks.[22] In other words, these scenarios play with the difference between $\hat{\boldsymbol{R}}_g$ and $\hat{\boldsymbol{R}}_r^0$, the estimated response of R to a contemporaneous innovation

[22]In Appendix A.1.5, Figure A.3 plots the impulse responses of R for each scenario.

to the monetary rule.

Figure 1.3 presents the simulation results. The left panel corresponds to the *misaligned* scenario, and the right panel corresponds to the *aligned* scenario. In each case, I show the estimated multipliers for four estimation strategies: (i) the TWFE approach, (ii) the control function approach, (iii) the ex-ante decomposition using only a contemporaneous innovation to R, (iv) the ex-ante decomposition using four news shocks[23] to R.

The *misaligned* scenario combines the strong forward-looking dynamics of the model (Del Negro et al., 2023; McKay et al., 2016) with significantly different paths for R after each aggregate shock. In this setting, the decomposition framework improves on the TWFE estimator, but suffers from some loss of precision. This is because the estimation is constrained by the limited number of shocks (i.e. $\tilde{F} = \{1, 4\}$). Nevertheless, the decomposition estimates, using one or four shocks, both outperform the control function approach, which in this calibration yields estimates that closely resemble those of the TWFE benchmark.

In the *aligned* scenario, the model still features forward-looking dynamics, but the interest rate responses to fiscal and monetary shocks are more closely aligned. Here, the decomposition performs substantially better. In particular, the ex-ante decomposition with one or four news shocks recovers a portable elasticity that is very close to the population value at all horizons. While the control function estimator also improves under this calibration, it continues to under-perform relative to the decomposition approach.

Extensions

Appendix A.1.5 presents several extensions of the simulation analysis. First, I show that the decomposition framework consistently recovers the portable elasticity as the number of available news shocks increases, under both parameterizations. Second, I examine environments with weaker forward-looking dynamics to confirm that the precision loss from limited information is smaller in those settings. Finally, in Subsection 1.5.3, I explore, within the limited-information

[23] That is, on top of the contemporaneous innovation, I assume that the researcher observes news shocks about the realization of R in period $t + s$ for $s = \{1, 2, 3\}$.

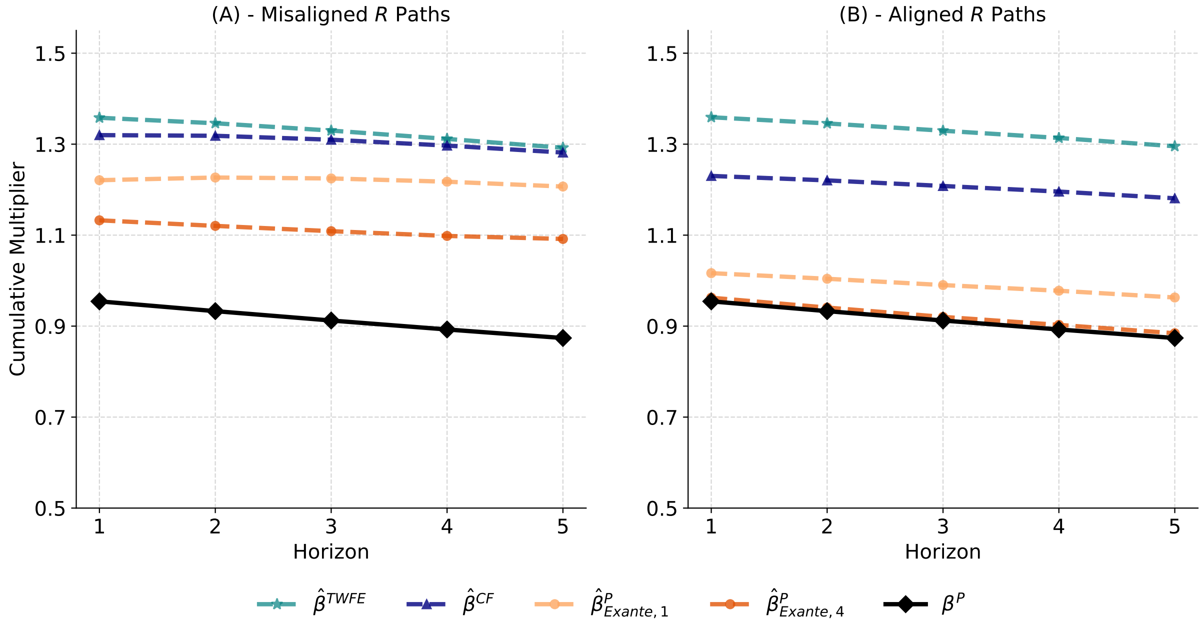


Figure 1.3. DSGE Monte Carlo Simulations - Results.

Note: The left (*right*) panel corresponds to a calibration where the path of R is *misaligned* (*aligned*) across aggregate shocks (See Figure A.3). The solid black line corresponds to β^P the population portable elasticities. The dashed lines correspond to different estimation methods. $\hat{\beta}^{TWFE}$: TWFE approach; $\hat{\beta}^{CF}$: control function approach; $\hat{\beta}^{P}_{Exante,k}$: ex-ante decomposition using k shocks.

setting, which observable data moments are informative about the accuracy of the framework and whether the method tends to over- or under-estimate the population portable elasticity, using the structure of the two-region New Keynesian model.

1.3.4 Asymptotic Properties and Inference

This section summarizes the asymptotic properties of the estimators and the corresponding inference procedure. A formal discussion of consistency is provided in Appendix A.1.6. The argument proceeds in three steps. First, consistency of the panel local projection estimators is established under large-N, fixed-T asymptotics. Second, consistency of the time-series local projection estimators is established under large-T asymptotics. Third, $\hat{\beta}^P_h$ is shown to be a continuous mapping of these estimators and therefore consistent by the Continuous Mapping Theorem under mild non-degeneracy conditions.

The estimated portable elasticities at horizon h depend on both panel and time-series coefficients across horizons. For inference, I jointly estimate the cross-sectional and time-series parameters using a just-identified GMM system, which facilitates inference across moment conditions. This approach allows me to explicitly account for the covariance between panel and time-series moment conditions induced by shared aggregate shocks by clustering at the time level.

For the covariance between panel moment conditions associated with different parameters, the framework can flexibly accommodate alternative clustering schemes (e.g., time, unit, or two-way clustering). For the time-series moments, I use a HAC-consistent estimator as the baseline.

Standard errors for the HEDGE terms and the portable elasticities are obtained via the Delta method applied to the cluster-robust covariance matrix of the moment conditions. Monte Carlo simulations show that this procedure delivers (i) standard errors equivalent to those from the control function approach in static settings, and (ii) correct coverage in dynamic environments. A detailed exposition of the inference procedure is provided in Appendix A.1.7.

Recap

This section outlined an estimation strategy to decompose cross-sectionally identified elasticities into a portable component and an HEDGE component. The framework combines cross-sectional and time-series analysis. In the cross-sectional step, it estimates the differential responses to innovations in the variable of interest, G, and to innovations in the variable(s) that the researcher wants to difference out, R, conditional on exposure to the main variable of interest s_{ig}. In the time-series step, it estimates the impulse response function of R to both the shock of interest, ε_{gt}, and to innovations to R itself, ε_{rt}. Finally, the output of these steps is combined to estimate the HEDGE component. I discuss how to apply each of these steps under different assumptions about the data-generating process and about the information available to the researcher. Using DSGE Monte Carlo simulations, I show that the decomposition framework identifies the portable elasticity. In cases where the researcher has access to limited information, the decomposition framework works as an approximation to the true portable elasticity, and it represents an improvement on both baseline estimates that ignore HEDGE and the correction implied by the control function approach.

1.4 Application - US Cross-sectional Fiscal Multiplier

To illustrate the use of the decomposition framework, I revisit the estimation of the US cross-sectional fiscal multiplier through the lens of an HEDGE economy. In particular, I evaluate the possibility that states in the US differ not only in how exposed they are to fiscal shocks but also to changes in interest rates. I focus on the link between government spending and interest rates for two reasons. First, the strength of the monetary policy response to government spending shocks (i.e., the degree of *monetary offset*) is an important determinant of the aggregate fiscal multiplier. The response of monetary policy is undoubtedly among the set of aggregate variables that we aim to difference out by using cross-sectional methods (Nakamura and Steinsson, 2014). Importantly, there is consensus in the profession that available estimates of the cross-sectional fiscal multiplier are independent of monetary policy, and the method proposed in this paper is designed to test whether this is indeed the case. Second, it is well documented that interest rate sensitivities vary substantially across economic regions (e.g. Carlino and DeFina (1998); Almgren et al. (2022); Herreño and Pedemonte (2025)), making it a natural starting point to explore the sensitivity of cross-sectional fiscal multipliers to HEDGE.

The empirical results should therefore be interpreted as isolating a single, yet economically central source of GE feedback—monetary policy. In this sense, the decomposition recovers a *monetary-policy–portable multiplier*, i.e., the cross-sectional effect of government spending holding the monetary response fixed. While other aggregate channels may also generate heterogeneous general equilibrium effects, I abstract from them in this application and leave their analysis to future work. Importantly, as shown below, removing this channel already restores the ability to use cross-sectional estimates to study counterfactual monetary policy regimes.

For estimation, I use the dataset built by Dupor and Guerrero (2017). The authors construct a panel dataset of US state-level defense contracts between 1951 and 2014.[24] Gross State Product data are sourced from BEA and available from 1963 onward. Following Sarto (2024), I exclude four small-population states with extreme variation in Gross State Product. Including all states yields similar point estimates with larger variance.[25] Both output and defense contracts are deflated using the national Consumer Price Index and scaled by state population. I complement this dataset with data on the Federal Funds rate, total nominal federal tax receipts, nominal GDP, and the series of Romer and Romer (2004) monetary policy shocks as extended by Wieland and Yang (2020). The frequency of analysis is yearly and covers the period 1969-2007.[26]

I construct output and defense spending variables to estimate cumulative fiscal multipliers following the recommendations of Ramey and Zubairy (2018). Concretely, the cumulative percentage change in variable X_{it} relative to a baseline level of output Y_{it-1} is computed as:

$$X^C_{it+h} = \frac{[\sum_{j=0}^{h} X_{it+j}] - (h+1)X_{it-1}}{Y_{it-1}} \times 100, \tag{1.74}$$

where the subscript i indexes US states and t indexes years. The aggregate variables are constructed analogously. I measure the state-level exposure to defense spending, s_{igt}, following Nakamura and Steinsson (2014):

$$s_{igt} = \frac{G^{pc}_{it}/Y^{pc}_{it}}{G^{pc}_{t}/Y^{pc}_{t}}, \tag{1.75}$$

where G^{pc}_{it} is per-capita real defense spending in state i during year t and Y^{pc}_{it} is per-capita real

[24]Data between 1951-2009 is sourced from two reports: the Prime Contract Awards by State report and the Atlas/Data Abstract for the US and Selected Areas. Both were compiled by the Directorate for Information Operations and Reports. Data for 2010-2014 are sourced from USASpending.gov.

[25]The excluded states are North Dakota (0.23% of the US population), South Dakota (0.27%), Wyoming (0.18%), and Alaska (0.23%). These states display multiple observations with GSP yearly changes over 20%. Figure A.18 in the Appendix presents results including all states.

[26]The sample is limited by the time-span of the Romer-Romer monetary policy shocks.

output in state i during year t. G_t^{pc} and Y_t^{pc} are the corresponding national counterparts. The idea is that a state is relatively more exposed to defense spending if its share of local defense spending to GDP is higher than the corresponding national share. I follow Nakamura and Steinsson (2014) and define s_{ig} as the average of this ratio over the first five years of their sample (i.e., 1966-1970).

First, I estimate the cross-sectional fiscal multiplier using the TWFE regression.

$$Y_{it+h}^{C} = \beta_h^{TW} \times G_{it+h}^{C} + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h}, \tag{1.76}$$

where Y_{it+h}^{C} is the cumulative change in real output per capita in state i, G_{it+h}^{C} is the cumulative change in per capita defense contracts in state i, and λ_{ih}, λ_{th} are horizon-specific fixed effects for each state and year. I follow the literature and instrument G_{it+h}^{C} with $s_{ig}G_t^{C}$, where s_{ig} denotes state i's exposure to aggregate defense spending. The instrument interacts this predetermined exposure with the contemporaneous aggregate change in defense contracts, G_t^{C}, thus isolating variation in regional spending driven by changes in aggregate defense spending. I control for the lag of real per-capita output growth. The parameter of interest is β_h^{TW}, which measures the response of local output to an increase in local government spending equivalent to 1% of local output - the cross-sectional fiscal multiplier. The superscript TW is used to denote that this estimate corresponds to the TWFE specification.

The standard practice in this literature is to cluster standard errors at the state (unit) level. However, as shown in Almuzara and Sancibrián (2024), the appropriate clustering dimension in shift-share settings with a common aggregate shock corresponds to the time dimension. Common aggregate shocks induce correlation in residuals across units within a given period. I conduct Monte Carlo simulations in Appendix A.1.7 that confirm this result: clustering at the unit level leads to substantial undercoverage at all horizons.

Accordingly, I report baseline results using time-level clustering for the panel estimates and follow the inference procedure outlined in Appendix A.1.7. This approach generally delivers larger standard errors than unit-level clustering.[27]

The left panel of Figure 1.8 plots the estimation results. The TWFE multiplier is 1.5 at the two-year horizon and above one at all shorter and longer horizons.

Next, I present evidence of HEDGE with respect to monetary policy responses and implement the decomposition framework as outlined in Section 1.3.2 using a sequence of counterfactual contemporaneous shocks and as outlined in Section 1.3.3 using only date *zero* shocks. In the following, I provide further details on each step.

1.4.1 HEDGE Test

I test for HEDGE in monetary policy responses by estimating the following local projections:

$$Y^C_{it+h} = c_h \times s_{ig} i_t + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h} \qquad \text{for} \quad h = 1, 4, \tag{1.77}$$

where i_t is the Federal Funds Rate. For estimation, I instrument $s_{ig} i_t$ with $s_{ig} RR_t$, where RR_t are the Romer-Romer monetary shocks.

Figure 1.4 shows that the estimated response is positive and statistically significant at all horizons. This finding rejects the null of homogeneous exposure to interest rate changes and suggests that time fixed effects alone may not fully absorb monetary policy shocks in this setting. In particular, the results indicate that states more exposed to defense spending, as measured by s_{ig}, are less responsive to changes in interest rates.

One possible explanation relates to differences in the cyclicality of government and private demand. If government purchases are less cyclical than private consumption, states with higher exposure to defense spending may be less sensitive to interest rate fluctuations. Another possibility is cross-state income differences: Figure A.10 in Appendix A.2 shows that per capita income is positively correlated with s_{ig}, and to the extent that higher income is associated with

[27] For comparison, Figure A.16 reports results using unit-level clustering.

weaker liquidity constraints, this pattern is consistent with existing evidence for the U.S. (Herreño and Pedemonte, 2025) and for Europe (Almgren et al., 2022).

Robustness

Appendix A.3.1 presents detailed results for each of the following robustness checks. First, in Figure A.12, I show that the estimated response is robust to including the interaction between government spending and state-level exposure $s_{ig}G_t^C$. Second, I re-estimate equation (1.77) using the monetary shock series from Aruoba and Drechsel (2024) for the period 1980-2007. Figure A.13 shows a positive estimated response, although less precisely estimated, possibly due to the shorter sample. Third, I estimate state-specific output elasticities to the Romer and Romer (2004) shock using local projections. Figure A.14, shows that the estimated state-level elasticities are positively correlated with state-level exposure to G and that the correlation is statistically significant.

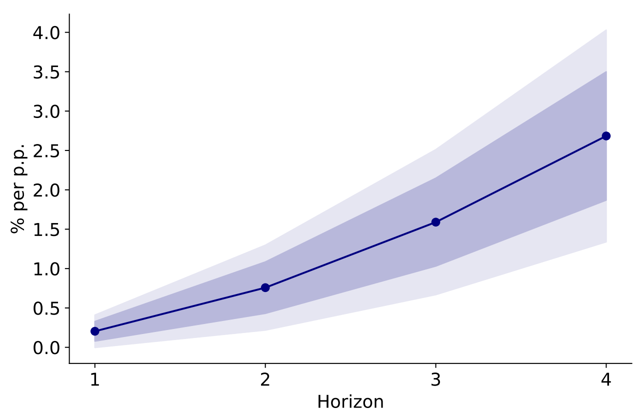


Figure 1.4. HEDGE Test - Response of Y_i to i_t, Conditional on s_{ig}.

Note: Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Plot corresponds to the cumulative percentage differential response of local output to a percentage point increase in the Federal Funds rate, evaluated at $s_{ig} = 1$.

The previous test indicates that in this empirical setting the cross-sectional condition for HEDGE to represent an identification challenge is met: s_{ig} drives cross-sectional responses to both defense and monetary policy shocks. Next, I show that the time-series condition is also met, namely, that the interest rate moves in response to a defense spending shock, which motivates the use of the decomposition framework.

1.4.2 Time-Series Step

The following step estimates the time series response of the interest rate to defense spending shocks and monetary policy shocks. First, this will inform whether the interest rate is effectively a GE variable in this empirical setting. I find this to be the case, so I use these reduced-form impulse responses to construct two different sequences of counterfactual monetary surprises. The first one, following Subsection 1.3.2, finds the sequence of ex-post contemporaneous surprises to the interest rate that replicate its observed path after a defense spending shock. I refer to this as the ex-post decomposition (Sims and Zha, 2006). Second, following Subsection 1.3.3, I find the magnitude of a period *zero* monetary surprise that better matches the estimated path of the interest rate after a defense spending shock in the least squares sense. I refer to this as the ex-ante decomposition (McKay and Wolf, 2023; Barnichon and Mesters, 2023).

First, I run the following local projections:

$$i_{t+h} = v_h G_t^C + a_h RR_t + \text{Controls} + e_{t+h}, \tag{1.78}$$

where i_t is the Federal Funds rate, G_t^C is the one-year change in national defense contracts relative to GDP and RR_t are the series of Romer and Romer (2004) monetary policy shocks. The set of controls includes lagged realizations of the dependent variable, the Romer-Romer shock, per-capita real output (in logs), the average federal tax rate, CPI inflation, and per-capita real defense spending (in logs). Standard errors are HAC robust.

Figure 1.5 shows the impulse response functions of the interest rate to each aggregate shock. In response to a defense spending shock equivalent to 1% of lagged national output, the

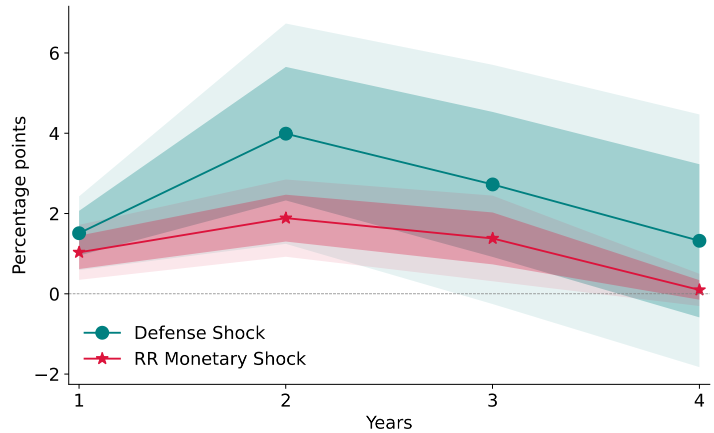


Figure 1.5. Time-series Step - Interest Rate Responses.

Note: Shaded area represents 68% (darker) and 90% (lighter) confidence bands. The green line shows the response of the nominal interest rate to a defense spending shock equal to 1% of lagged national output. The red line shows the response to a Romer–Romer monetary policy shock of one standard deviation.

interest rate increases by around 1.5 percentage points on impact.[28] A one-standard-deviation Romer–Romer shock raises the interest rate by 1 percentage point on impact, with a peak response of approximately 1.9 percentage points in the second year. Importantly, the shape of these two impulse responses is relatively similar, which, as I discuss in Subsection 1.5.3, indicates that the loss of precision in applying the decomposition without access to news shocks should be relatively small.

In order to implement the ex-post decomposition, I employ the estimated dynamic response to the Romer-Romer monetary shock to compute up to $H = 4$ counterfactual monetary surprises that replicate the interest rate response to the defense spending shock. Figure 1.6A displays the resulting sequence of shocks, denoted $\tilde{\boldsymbol{e}}$.

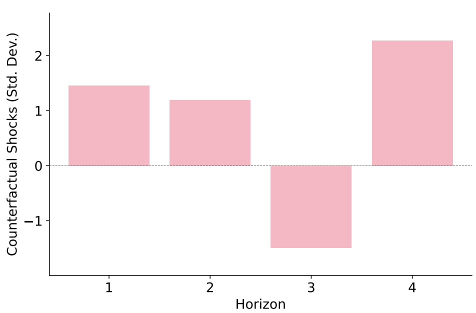
For the ex-ante decomposition, I only use a single counterfactual shock $\tilde{\boldsymbol{e}}^0$ that hits the economy at $h = 0$ and propagates according to $\hat{a}_h$. This involves solving the minimization problem (1.79) to find the size of the Romer–Romer shock that best replicates the path of the

[28]The response to a one standard deviation increase in G_t^C is of approximately 0.42 percentage points.

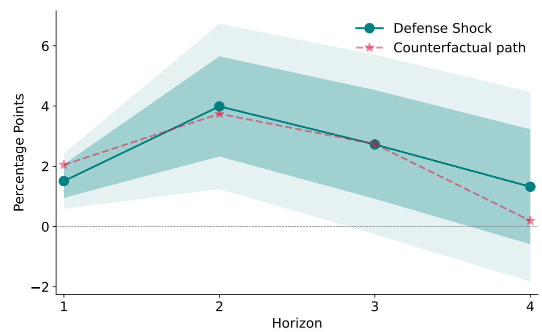
interest rate following the defense shock. Setting $F = 0$, the minimization problem becomes:

$$\tilde{\boldsymbol{e}}^0 = \arg\min_{e \in \mathbb{R}} \sum_{h=0}^{H} (\hat{v}_h - \hat{a}_{h0}\, e)^2, \tag{1.79}$$

where $\hat{a}_{h0} = \hat{a}_h$ are the estimated responses to a Romer-Romer shock. I find that a Romer-Romer shock of $\tilde{\boldsymbol{e}}^0 = 1.98$ standard deviations minimizes (1.79). Figure 1.6B plots the resulting ex-ante counterfactual path for the interest rate, alongside its estimated response to the defense shock. The counterfactual path remains within one standard deviation of the target response at all horizons.



(A) Ex-post Counterfactual R Shocks.



(B) Ex-ante Counterfactual Path of R.

Figure 1.6. Time Series Step - Counterfactual Computations.

Note: Panel (A) shows the sequence of estimated counterfactual monetary shocks for the ex-post decomposition. The units are standard deviations of Romer-Romer monetary shocks. Panel (B) shows the counterfactual path for R implemented by the ex-ante decomposition with $\tilde{\boldsymbol{e}}^0 = 1.98$. *Defense Shock* is the estimated response of the interest rate to the defense shock with its corresponding 68% and 90% confidence bands.

1.4.3 Cross-sectional Step

I jointly estimate the cross-sectional responses to defense spending shocks and interest rate shocks using the following reduced-form specification:

$$Y^C_{it+h} = b_h \times s_{ig} G^C_t + c_h \times s_{ig} RR_t + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h}, \tag{1.80}$$

where RR_t is the series of Romer and Romer (2004) monetary shocks and G^C_t is the one-year change in aggregate defense contracts. I include the same set of controls as in the TWFE specification. The coefficient c_h captures the differential response of cumulative output to a one standard deviation in the monetary shock for the region with exposure equal to the national average. I separately estimate the first-stage regression for local government spending using:

$$G^C_{it+h} = \theta_h \times s_{ig} G^C_t + \lambda_{ih} + \lambda_{th} + Controls + e_{it+h}, \tag{1.81}$$

where θ_h captures the first-stage coefficient at horizon h. I add as controls all variables included in the reduced form specification (1.80). I proceed in two steps to align the units from the time-series and cross-sectional reduced-form estimates.

Figure 1.7 shows the percentage response of local output to a monetary shock of one standard deviation, evaluated at $s_{ig} = 1$. The estimated differential response is positive and significant at all horizons, consistent with the findings in the HEDGE test above. The corresponding two-stage least squares responses to the defense shock, $\hat{b}_h / \hat{\theta}_h$, are nearly identical to the TWFE estimates and are therefore relegated to Figure A.15 in Appendix A.3.2.[29]

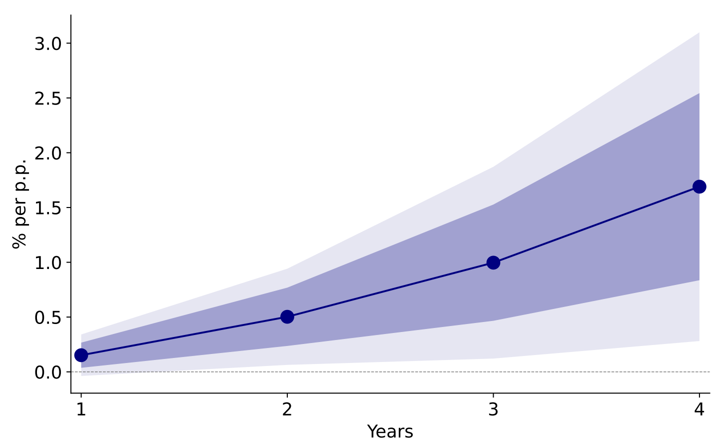


Figure 1.7. Response of Y_i to RR_t, Conditional on s_{ig}.

Note: Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.1.7 with time-level clustering for panel moment conditions. The solid line shows the cumulative percentage response of local output to a one standard deviation Romer-Romer shock, evaluated at $s_{ig} = 1$. The specification controls for lagged output growth.

[29] These estimates may differ from $\hat{\beta}^{TW}_h$ if there is in-sample correlation between the Romer-Romer shock and any of the other regressors included in the specification.

1.4.4 Portable Multiplier Estimate

Lastly, I combine the results of the previous two steps to construct estimates for the portable cross-sectional multiplier up to $H = 4$. Formally, $\hat{\beta}_h^P$ is given by:

$$\hat{\beta}^P_{h,Expost} = \frac{\hat{b}_h - \sum_{k=0}^{h} \hat{c}_{h-k} \times \tilde{e}_{r,t+k}}{\hat{\theta}_h}, \qquad \hat{\beta}^P_{h,Exante} = \frac{\hat{b}_h - \hat{c}_h \times \tilde{\boldsymbol{e}}^0}{\hat{\theta}_h}, \tag{1.82}$$

where $\tilde{e}_{r,t+k}$ is the ex-post monetary surprise in period $t+k$ and $\tilde{\boldsymbol{e}}^0$ is the period *zero* monetary shock. Dividing by the first stage coefficient expresses results in terms of changes to local government spending as is usual in this literature. Figure 1.8 shows the results for each version of the decomposition and for the TWFE specification.[30] The control function approach, shown for completeness, delivers estimates that are virtually identical to those of the TWFE strategy. The portable multiplier is below the TWFE multiplier at all horizons for both versions. This is for two reasons. First, it reflects the fact that $\hat{c}_h > 0$ which means that US states that are more exposed to defense spending also react less to changes in interest rates. Second, the interest rate is estimated to increase in response to a defense spending shock. Together, both forces push the cross-sectional multipliers estimated using the TWFE strategy above the estimated portable multipliers. For example, at the two-year horizon, the TWFE multiplier is 1.5 while the estimated portable multiplier is 1.

Lastly, Figure 1.9 shows the estimated HEDGE terms for each version of the decomposition expressed in the same units as the cross-sectional multiplier:

$$\tilde{\Omega}_{h,Expost} = \frac{\hat{\Omega}_{h,Expost}}{\hat{\theta}_h}, \qquad \tilde{\Omega}_{h,Exante} = \frac{\hat{\Omega}_{h,Exante}}{\hat{\theta}_h}, \tag{1.83}$$

where $\hat{\theta}_h$ are the first stage coefficients from (1.81). The estimated HEDGE terms are statistically different from zero at the 90% level for $h > 1$ for both versions of the decomposition. For comparison, Figure A.17 reports results clustering at the unit-level.

[30] See Figure A.16 for unit-level clustering.

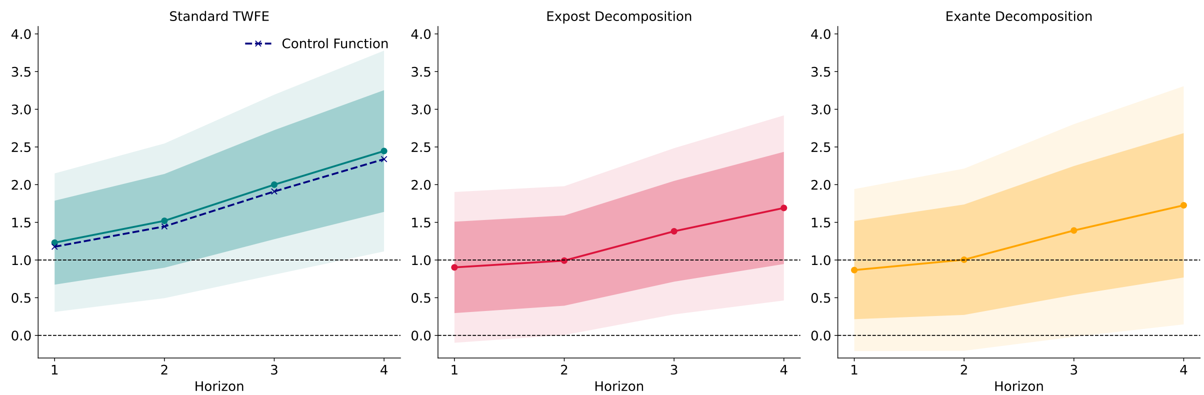


Figure 1.8. Decomposition Results for the US Cross-Sectional Multiplier.

Note: Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.1.7 with time-level clustering for panel moment conditions. Point estimates show the cumulative percentage change in local output when local defense spending increases by 1% of local output. *Ex-post Decomposition* shows results for the decomposition outlined in Subsection 1.3.2. *Ex-ante Decomposition* shows results for the decomposition outlined in Subsection 1.3.3 using a single shock. The blue dotted line shows the results from the control function approach.

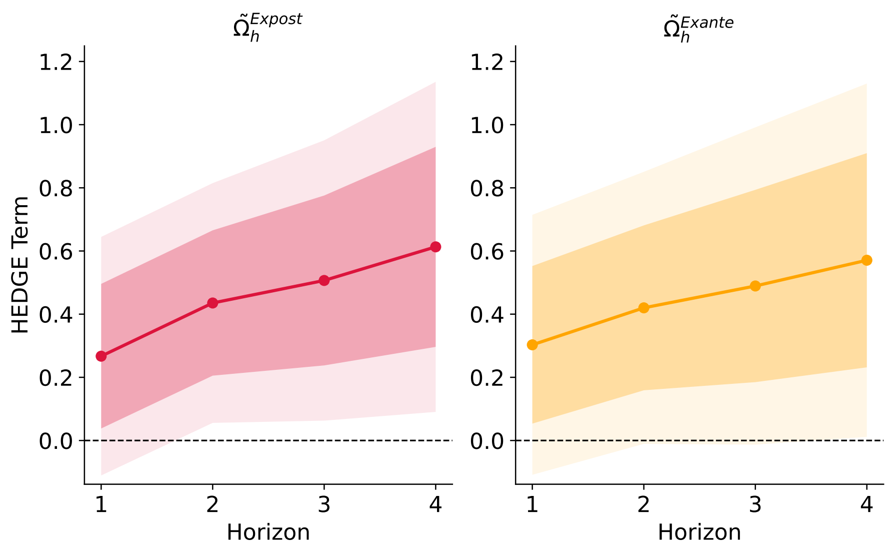


Figure 1.9. Decomposition Results - Estimated HEDGE Terms.

Note: The plot shows the point estimate and confidence bands for $\tilde{\Omega}_{h,Expost}$ on the left-hand side and for $\tilde{\Omega}_{h,Exante}$ on the right-hand side. Standard errors are computed as described in Appendix A.1.7 with time-level clustering for panel moment conditions.

Robustness

Appendix A.3.2 presents results from a series of robustness checks. First, Figure A.16 and Figure A.17 report the baseline specification using unit-level clustering for panel moments instead of time-level clustering. As expected, this choice yields narrower confidence bands. Second, I show that results are robust to using the real Federal Funds rate instead of the nominal rate. Figure A.19 shows that the point estimates for the portable multiplier remain largely unchanged, although the ex-post decomposition becomes noisier. Third, Figure A.20 shows that the results are robust to using the period average value of s_{igt} instead of the average of 1966-1970. Fourth I present results using two alternative sets of controls in the cross-sectional regressions: (i) no controls, (ii) baseline plus lagged shock(s). Figure A.21 shows results when no controls are included - in line with the baseline choices of, for example, Nakamura and Steinsson (2014); Dupor and Guerrero (2017); Auerbach et al. (2020). This choice results in significantly larger estimates of the cross-sectional fiscal multiplier relative to my baseline specification, which controls for lagged output growth. The fact that controlling for lagged output dampens the cross-sectional multiplier was first noted by Ramey (2020). Overall, both the TWFE and portable multipliers shift up, yet the magnitude of the estimated HEDGE term remains stable. Figure A.22 corresponds to the baseline specification augmented by adding lags of the corresponding aggregate shocks, interacted with s_{ig}. Adding lagged shocks further shifts downward both the TWFE and portable estimated multipliers, in particular, both drop below 1 at the two-year horizon. However, the estimated HEDGE term is of the same magnitude as in the baseline specification, albeit much less precisely estimated. Fifth, I present the results for seven alternative control choices detailed in Table A.8. Figure A.23A plots the baseline estimate for the TWFE alongside each of the seven alternatives. Figures A.23B and A.23C repeat this exercise for the ex-ante portable multiplier and the HEDGE term, respectively. Across specifications, the estimated HEDGE terms fall within one-standard deviation of the baseline estimates.

Recap

This section revisits the estimation of cross-sectional fiscal multipliers using state-level data for the US between 1969-2007. I show that taking into account the heterogeneous effects of monetary policy across US states matters for the estimation of cross-sectional fiscal multipliers.

In particular, I estimate a portable multiplier that is meaningfully lower than baseline TWFE estimates. This finding underscores the limitations of relying solely on time fixed effects to difference-out GE effects, in particular, those operating through systematic monetary responses.

1.5 Model

This section develops a theoretical counterpart to the empirical decomposition of Section 1.4 using a two-region New Keynesian environment. The model provides a structural interpretation of the empirical objects by expressing the TWFE cross-sectional multiplier—denoted by $\mathscr{M}_{CS}$—the HEDGE term Ω, and the portable multiplier—denoted by $\mathscr{M}_{CS}^{P}$—in terms of model-implied responses to a fiscal policy shock. I use the model to decompose $\mathscr{M}_{CS}$ into three channels: substitution between regional goods; intertemporal substitution driven by regional inflation differentials; and the HEDGE channel, which arises when regions differ in their responsiveness to aggregate real interest rate movements. The HEDGE term enters as a residual general-equilibrium force that moves the model-analogue of the TWFE multiplier away from its portable counterpart.

Next, I show that HEDGE can cause cross-sectional multipliers to vary widely across monetary regimes, sometimes in directions opposite to those of the aggregate multiplier. This challenges the common interpretation of $\mathscr{M}_{CS}$, the TWFE cross-sectional multiplier, as partial equilibrium objects, a view often justified by the inclusion of time fixed effects. Crucially, when regions differ in their exposure to general-equilibrium forces, $\mathscr{M}_{CS}$ reflects not only direct policy effects, but also heterogeneous responses to GE adjustments. As a result, the property highlighted in prior work (Chodorow-Reich, 2019; Dupor et al., 2023) that cross-sectional multipliers can serve as upper or lower bounds on the aggregate multiplier, depending on the monetary environment, is no longer guaranteed when HEDGE is present and not taken into account. In contrast, the portable multiplier $\mathscr{M}_{CS}^{P}$ - for which this paper provides an identification strategy - retains several key properties: GE invariance, validity as a bound on the aggregate multiplier, and usefulness for model comparison.

Finally, I treat the two-region New Keynesian model as a useful approximation to the empirical environment and ask: which observable data moments best predict the precision loss in estimating the HEDGE term under limited information? The model reveals that a sufficient statistic, the cumulative discrepancy between the realized and counterfactual path of R, is highly informative about the estimation error in the HEDGE term. I then calibrate the model to match the empirical value of this moment. This allows me to evaluate how the decomposition performs in a data-generating process that mirrors the empirical setting in Section 1.4. The results suggest that, conditional on the model environment, the method delivers accurate estimates of the portable multiplier even in the presence of strong forward-looking dynamics.

Environment

I now present a general model structure, beginning with its sequence-space representation and structural blocks. The derivations follow McKay and Wolf (2023); Wolf (2023a), and the two-region[31] model is based on Nakamura and Steinsson (2014). The economy consists of two regions and is characterized by a set of $n_k \times 2$ endogenous regional variables $\{\mathbf{K}_t\}$ and a set of n_w endogenous aggregate variables $\{\mathbf{W}_t\}$. There are two structural aggregate shocks $s \in \{w_1, w_2\}$ that represent changes in aggregate variables $w_{1t}, w_{2t} \in \mathbf{W}_t$. The time path of a shock s is denoted by ε_s, and $\boldsymbol{\varepsilon} = (\varepsilon_{w_1}, \varepsilon_{w_2})$ denotes the full path of the shock. Each entry j in ε_s corresponds to an innovation in w_{st} that is announced at time $t = 0$ and implemented at time $t = j$.[32] Let $d\mathbf{x} = \{x_t\}_{t=0}^{\infty}$ denote the perfect-foresight path of a variable x, and let $d\mathbf{x}_{|\boldsymbol{\varepsilon}}$ be the corresponding path conditional on the shock sequence $\boldsymbol{\varepsilon}$. The regional block of the model is characterized by

$$\mathcal{H}_k d\mathbf{K} + \mathcal{H}_w d\mathbf{W} = \mathbf{0}, \tag{1.84}$$

[31]The framework can be generalized to arbitrary N regions.

[32]For example, if ε_i is the vector of innovations to the nominal interest rate, then the first entry captures a contemporaneous surprise, while the second represents news received at $t = 0$ about a change in the rate at $t = 1$.

where the vector $d\mathbf{K}$ collects the time paths of $k_r \in \mathbf{K}_t$ and similarly for $d\mathbf{W}$. The matrices $\mathcal{H}_k$ and $\mathcal{H}_w$ condense the relationships that characterize the regional block.[33] The aggregate block is characterized by

$$\mathcal{A}_k d\mathbf{K} + \mathcal{A}_w d\mathbf{W} + A_\varepsilon \boldsymbol{\varepsilon} = \mathbf{0}. \tag{1.85}$$

The above formulation assumes that the structural shocks *only* affect the regional block through their effect on the set of aggregate variables $\mathbf{W}_t$. This assumption is important because it ensures that these shocks serve as valid instruments for the observed realizations of the policy or aggregate variables.

I denote by $\boldsymbol{H}_z^x$ the partial equilibrium Jacobian of z with respect to x. This matrix maps changes in the time-path of variable x to changes in variable z, holding all else equal. Each column j of this matrix gives the path of z in response to a one-unit innovation to x announced at $t = 0$ and implemented at $t = j$. The general equilibrium Jacobian of z with respect to the structural shock ε_{w_1} is denoted by $\boldsymbol{J}_z^{\varepsilon_{w_1}}$. Each column j of this matrix contains the general equilibrium response of an endogenous variable z to a change in w_1 announced in period $t = 0$ and implemented in period $t = j$.

1.5.1 Anatomy of the NK Cross-sectional Fiscal Multiplier

I use these objects to characterize the determinants of the cross-sectional fiscal multiplier in a standard two-region New Keynesian model. I take as a baseline the model in Nakamura and Steinsson (2014) and extend it to allow regional heterogeneity in household responses to the real interest rate. While the full model features both Ricardian and hand-to-mouth households, the derivations below assume a representative household in each region for tractability. The main features of the model are summarized here, with full details presented in the Appendix A.4.

The economy consists of two regions of equal size n, denoted H (Home) and F (Foreign).

[33] I omit regional shocks and unobserved variables for notational simplicity but the environment can be generalized to accommodate those features.

In each region, a representative household has separable preferences over labor and consumption and CES preferences over domestically and foreign-produced goods, each of which is itself a CES aggregate of locally produced varieties. Financial markets are complete, which implies perfect risk sharing across regions.

The federal government purchases domestic and foreign varieties using the same CES demand structure of households. Aggregate government spending is subject to an idiosyncratic shock ε_g that increases spending heterogeneously across space: each additional dollar of spending increases purchases of the Home (Foreign) good by $s_H \times n$ ($s_F \times n$) dollars. The government levies lump sum taxes to ensure the budget balances in each period.

On the supply side, local firms produce differentiated varieties using local labor as the sole input. Firms compete monopolistically and set prices à la Calvo, giving rise to two regional New Keynesian Phillips curves. Lastly, the nominal interest rate is set by a monetary authority and is subject to an idiosyncratic shock ε_i.

Aggregate government spending, G, follows an exogenous process governed by:

$$d\boldsymbol{G} = [\boldsymbol{A}_G^G]^{-1} \varepsilon_g, \tag{1.86}$$

where $\boldsymbol{A}_G^G$ allows for a general auto-regressive structure. This implies the following process for regional government consumption g_r:

$$\begin{aligned} d\boldsymbol{g_r} &= \boldsymbol{H}_{g_r}^G d\boldsymbol{G} \\ &= s_r \, d\boldsymbol{G} \qquad \forall \; r = H, F. \end{aligned} \tag{1.87}$$

The scalar s_r captures the exposure of the region r to aggregate government purchases. The nominal interest rate is set according to the following inertial Taylor rule:

$$d\boldsymbol{i} = \boldsymbol{A}_i^Y d\boldsymbol{Y} + \boldsymbol{A}_i^\Pi d\boldsymbol{\Pi} + \boldsymbol{A}_i^{\varepsilon_i} \varepsilon_i, \tag{1.88}$$

where Y and Π denote aggregate output and national inflation, respectively (i.e., the weighted sum of the corresponding regional counterparts). Total consumption in region r is determined by the Euler equation:

$$d\boldsymbol{c_r} = \boldsymbol{H}^r_{c_r} d\boldsymbol{r} + \boldsymbol{H}^q_{c_r} d\boldsymbol{q} \quad \forall \ r = H, F. \tag{1.89}$$

Consumption depends on the path of the national real interest rate, $r \in \boldsymbol{W}$, and the real exchange rate, $q \in \boldsymbol{K}$, which captures regional inflation differentials. The real exchange rate reflects the gap between the national and local real rates induced by differences in regional CPI inflation.

The path of local output can be expressed as a function of regional demand, regional government purchases, and relative prices across regions.

$$d\boldsymbol{Y_r} = \boldsymbol{H}^{c_H}_{y_r} d\boldsymbol{c_H} + \boldsymbol{H}^{c_F}_{y_r} d\boldsymbol{c_F} + \boldsymbol{H}^{g_r}_{y_r} d\boldsymbol{g_r} + \boldsymbol{H}^q_{y_r} d\boldsymbol{q} \quad \forall \ r = H, F. \tag{1.90}$$

The first two terms on the right-hand side capture the effects of changes in regional consumption demand. The third term captures the effect of regional government purchases on local output. The last term measures the effect of changes in q, the real exchange rate between regions.

Let $\boldsymbol{\varepsilon}_g$ denote a time path of structural shocks such that $\varepsilon_g = [1, 0, ..0]$ and $\varepsilon_i = \boldsymbol{0}$. In other words, there is a contemporaneous one-time increase in ε_g that propagates according to (1.86) and no further shocks hit the economy. We want to compute the cross-sectional output multiplier, $\mathscr{M}_{CS}$, in response to $\boldsymbol{\varepsilon}_g$:

$$\mathscr{M}_{CS} = \frac{d\boldsymbol{y}_{\boldsymbol{H}|\varepsilon_g} - d\boldsymbol{y}_{\boldsymbol{F}|\varepsilon_g}}{d\boldsymbol{g}_{\boldsymbol{H}|\varepsilon_g} - d\boldsymbol{g}_{\boldsymbol{F}|\varepsilon_g}}. \tag{1.91}$$

where division between vectors is understood element-wise across horizons. Since $d\boldsymbol{g}_r = s_r d\boldsymbol{G}$, we have:

$$dg_H - dg_F = (s_H - s_F)d\boldsymbol{G}, \tag{1.92}$$

which yields

$$\mathscr{M}_{CS} = \frac{d\boldsymbol{y}_{\boldsymbol{H}|\varepsilon_g} - d\boldsymbol{y}_{\boldsymbol{F}|\varepsilon_g}}{(s_H - s_F)\ d\boldsymbol{G}_{|\varepsilon_g}}, \tag{1.93}$$

where $\mathscr{M}_{CS}$ measures the relative change in local output per unit of differential government spending between regions after an aggregate fiscal shock.[34] Substituting in the structural expressions for local output and consumption yields a decomposition into three components: two relative price effects and one interest rate sensitivity term.

$$\mathscr{M}_{CS} = d\mathbf{1} + \frac{1}{(s_H - s_F)\ d\boldsymbol{G}_{|\varepsilon_g}}\left[(\Theta_{IS} + \Theta_{ES})d\boldsymbol{q}_{|\varepsilon_g} + \vartheta[\boldsymbol{H}^r_{c_H} - \boldsymbol{H}^r_{c_F}]d\boldsymbol{r}_{|\varepsilon_g}\right], \tag{1.94}$$

where $d\mathbf{1}$ is a vector of ones with the same dimension as $d\boldsymbol{G}$. In this model, the cross-sectional fiscal multiplier can deviate from unity through three channels.[35]

First, an *expenditure switching* channel, which captures how households reallocate consumption between Home and Foreign goods in response to changes in relative prices. Its strength is governed by Θ_{ES}, a function of the elasticity of substitution across Home and Foreign goods and the degree of home bias in preferences. A positive fiscal shock raises the relative price of goods produced in the more exposed region, leading households to substitute toward the less exposed region. This substitution dampens the local output response, reducing the multiplier.

The second is a *relative intertemporal substitution* channel that reflects the effect of differences in local real rates on consumption. These differences arise due to heterogeneous

[34]Without loss of generality, the above formula can be modified to compute either cumulative multipliers or multipliers relative to the change in government purchases on impact.

[35]In the full model, the presence of *hand-to-mouth* households adds an additional *Keynesian* channel that, generally, amplifies the cross-sectional multiplier.

consumption baskets across regions - for example, from home-biased preferences - which cause region-specific CPI inflation responses. This channel operates through movements in the real exchange rate and its strength is governed by Θ_{IS}. Under home bias, the relative intertemporal channel also has a negative effect on the size of the multiplier. The fiscal shock generates expected relative disinflation in the more exposed region, raising its relative real interest rate and lowering its relative consumption.[36] Like the ES channel, this channel tends to reduce the multiplier.

The third channel, which I refer to as the *HEDGE* channel, also operates through intertemporal substitution but captures differences in sensitivity to the national real rate between regions. This channel is non-zero only when regions differ in their interest rate sensitivity. For example, with CRRA utility and separable preferences, this would be the case with region-specific intertemporal elasticities of substitution (IES). Its strength is governed by two components: (i) ϑ, a function of the degree of home bias in preferences; and (ii) the heterogeneity in real rate sensitivity. Unlike the other two channels, HEDGE can amplify or dampen the cross-sectional multiplier, depending on the direction of the real rate movement and the correlation between exposure to fiscal policy and interest rate sensitivity.

The portable multiplier is defined as the TWFE cross-sectional multiplier net of the HEDGE channel:

$$\mathscr{M}^{P}_{CS} = \mathscr{M}_{CS} - \underbrace{\frac{\vartheta[\boldsymbol{H}^{r}_{c_H} - \boldsymbol{H}^{r}_{c_F}]}{(s_H - s_F)} \frac{d\boldsymbol{r}_{|\boldsymbol{\varepsilon}_g}}{d\boldsymbol{G}_{|\boldsymbol{\varepsilon}_g}}}_{\text{HEDGE Channel}}. \tag{1.95}$$

This mirrors the decomposition in equation (1.2) and identifies the model-based analogue to the empirical HEDGE term, Ω_h, and to the portable elasticity, β^{P}_{h}. The first term in the HEDGE channel captures the relative sensitivity to the real rate, normalized by fiscal exposure. The

[36]On impact, the spending shock increases relative producer price inflation for the exposed region. Regional convergence requires that this initial difference in inflation rates across regions be compensated for by expected relative disinflation in the future. That is, there is a decrease in expected future producer price inflation in the exposed region relative to the less exposed region. With home bias, this translates into a decrease in expected consumer price inflation for the exposed region because its consumption basket is loads more heavily on its locally produced good.

second captures the general equilibrium response of the real rate to government spending shocks.

To compute the model analogue of the empirical HEDGE term Ω_h, I isolate the contribution of the real interest rate channel to the differential in cross-regional output $d\boldsymbol{y}_{\boldsymbol{H}|\boldsymbol{\varepsilon}_g} - d\boldsymbol{y}_{\boldsymbol{F}|\boldsymbol{\varepsilon}_g}$. I ask: What portion of the differential would arise if the only effect of the fiscal shock were the general equilibrium path of the real rate, holding all other channels constant? To answer this, I construct a counterfactual sequence of monetary policy shocks that reproduces the real rate path $d\boldsymbol{r}_{|\boldsymbol{\varepsilon}_g}$ induced by the original fiscal shock. This allows me to simulate the output effects of the HEDGE channel in isolation and recover the model-implied path of Ω.

Under the assumption that structural shocks affect regional outcomes only through their impact on aggregate variables, the HEDGE term can be computed using the general equilibrium response of local output to a time-path of monetary innovations $\tilde{\boldsymbol{\varepsilon}}_i$ that replicates $d\boldsymbol{r}_{|\boldsymbol{\varepsilon}_g}$. Letting $\boldsymbol{G}_r^{\varepsilon_i}$ denote the general equilibrium Jacobian of the real rate with respect to the monetary shock ε_i, I compute the following:

$$\tilde{\boldsymbol{\varepsilon}}_i = [\boldsymbol{G}_r^{\varepsilon_i}]^{-1} d\boldsymbol{r}_{|\boldsymbol{\varepsilon}_g}, \tag{1.96}$$

where $\tilde{\boldsymbol{\varepsilon}}_i$ is the vector of contemporaneous and news shocks that implements the desired real rate path. Given this sequence, the region-specific output responses are:

$$d\boldsymbol{y}_{r|\tilde{\boldsymbol{\varepsilon}}_i} = \boldsymbol{G}_{y_r}^{\varepsilon_i} \tilde{\boldsymbol{\varepsilon}}_i \quad \forall\, r = H, F. \tag{1.97}$$

The implied model counterpart to Ω_h is given by:

$$\boldsymbol{\Omega} = \frac{d\boldsymbol{y}_{H|\tilde{\boldsymbol{\varepsilon}}_i} - d\boldsymbol{y}_{F|\tilde{\boldsymbol{\varepsilon}}_i}}{(s_H - s_F)\, d\boldsymbol{G}_{|\boldsymbol{\varepsilon}_g}}, \tag{1.98}$$

where $\boldsymbol{\Omega} = \{\Omega_h\}_{h=0}^{H}$ denotes the vector of model-implied HEDGE terms across horizons. Subtracting this vector from the TWFE multiplier $\mathscr{M}_{CS}$ yields $\mathscr{M}_{CS}^{P}$, the model-analogue of the portable cross-sectional multiplier.

1.5.2 Pitfalls of Failing to Control for HEDGE

The previous discussion highlighted how the HEDGE channel may dampen or amplify the cross-sectional fiscal multiplier relative to a counterfactual economy where regions are equally responsive to interest rate changes. In this section, I use the model to study the implications of failing to control for HEDGE on (i) the dispersion of cross-sectional fiscal multipliers and on (ii) their relation to the aggregate fiscal multiplier - denoted by $\mathscr{M}_A$.

The first point relates to the monetary policy invariance that has been attributed to the available estimates of the cross-sectional multiplier. In HEDGE economies, the New Keynesian model can generate cross-sectional fiscal multipliers $\mathscr{M}_{CS}$, which are just as dispersed as the aggregate multipliers.

The exercise is as follows. Consider three different monetary policy rules that have different implications for the covariance between government spending shocks and real interest rates. These are the same three rules studied by Nakamura and Steinsson (2014). In the first case, the monetary authority follows an inertial Taylor rule with a weight on inflation of $\phi_\pi = 1.5$ and on the output gap of $\phi_y = .1$. This monetary rule implies an aggressive response to the inflationary pressures that follow a government spending shock. For the second case, I assume that the monetary authority keeps the national real rate fixed at its steady-state level in response to government spending shocks (i.e. $i_t - \mathbb{E}_t\left[\Pi_{t+1}^{agg}\right] = r_{ss} = i_{ss}$). In the third case, the monetary authority keeps the nominal interest rate fixed at its steady-state level in response to government spending shocks (i.e. $i_t = i_{ss}$). This last scenario is akin to an economy that hits the Zero Lower Bound (ZLB).

In addition, I consider two scenarios for the covariance between exposure to government spending shocks and to interest rate changes. In both cases, the shock is to ε_g but the difference comes from how the aggregate shock loads in each region. I refer to the case where the more interest rate sensitive region is more exposed to government spending as *high IES*, and to the reverse case as *low IES*. Essentially, I solve the model for each combination of monetary rule and spending shock structure to compute the corresponding cross-sectional and aggregate fiscal multipliers. The calibration is the one used for the DSGE Monte Carlo exercise in Subsection 1.3.3, detailed in Table A.9.

Table 1.2. Fiscal Multipliers & Monetary Policy with HEDGE.

	Government Spending Shock	
Monetary Rules	*High IES*	*Low IES*
Taylor Rule		
Portable Cross Sectional Multiplier - $\mathscr{M}_{CS}^{P}$	.95	.95
TWFE Cross Sectional Multiplier - $\mathscr{M}_{CS}$	.6	1.3
Aggregate Multiplier - $\mathscr{M}_{A}$	.64	.64
Fixed Nominal Rate		
Portable Cross Sectional Multiplier - $\mathscr{M}_{CS}^{P}$	.95	.95
TWFE Cross Sectional Multiplier - $\mathscr{M}_{CS}$	1.07	.82
Aggregate Multiplier - $\mathscr{M}_{A}$	1.12	1.14
Fixed Real Rate		
Portable Cross Sectional Multiplier - $\mathscr{M}_{CS}^{P}$	.95	.95
TWFE Cross Sectional Multiplier - $\mathscr{M}_{CS}$	.95	.95
Aggregate Multiplier - $\mathscr{M}_{A}$	1	1

Note: The table shows one year multipliers for alternative monetary rules and cross-sectional loadings of the government spending shock.

Table 1.2 presents the model-analogues of the TWFE and portable cross-sectional multipliers, as well as the aggregate multiplier for each scenario. The portable cross-sectional multiplier is, by definition, invariant to monetary policy and to the structure of the government spending shock. However, the TWFE cross-sectional multiplier depends on both the monetary response and on which region is more exposed to the spending shock. These multipliers range between .6 and 1.3, whereas the aggregate multiplier, which also depends on the monetary rule, ranges between .64 and 1.14. This exercise illustrates that HEDGE can introduce substantial variability in cross-sectional multipliers depending on the direction of GE responses and the correlation between cross-sectional exposures.

The results in Table 1.2 also show how general equilibrium forces that attenuate the aggregate effect of a fiscal shock, like the aggressive monetary response implied by the Taylor rule case, can actually increase the estimated cross-sectional effect, conditional on the structure of the government spending shock. For example, a researcher studying the effect of a government spending shock tilted to the *low IES* region in a ZLB environment would estimate a smaller cross-sectional multiplier than a researcher doing so under normal monetary policy. The opposite is true with a shock tilted to the *high IES* region. The counterintuitive and context-specific way in which HEDGE affects cross-sectional estimates brings us to the second point of this section. Namely, how informative are TWFE cross-sectional estimates of the aggregate fiscal multiplier?

Previous work has argued that under conventional monetary policy, such as a Taylor rule that responds aggressively to inflation, the cross-sectional multiplier is likely to serve as an upper bound on the aggregate multiplier. The reason is that TWFE cross-sectional estimators do not include the contractionary effects of monetary tightening, which typically dominate the dampening effects of expenditure switching or relative intertemporal substitution. Conversely, at the ZLB, monetary policy is unresponsive, so cross-sectional estimates tend to understate the aggregate effects, acting as a lower bound. Table 1.2 shows that these conclusions may no longer hold once HEDGE is present and is not accounted for. In particular, changes in the bias of the government spending shock or the differential interest rate sensitivities across regions can flip these bounding relationships in either direction.

This highlights a key point: When general equilibrium responses to monetary policy contaminate cross-sectional estimates, the sign and magnitude of the resulting bias become context-specific. As a result, TWFE cross-sectional multipliers cannot be reliably interpreted as bounds on aggregate effects. In contrast, the portable multiplier $\mathscr{M}^{P}_{CS}$, which I show how to identify empirically, retains its validity to bound the aggregate fiscal multiplier regardless of the monetary policy regime.

1.5.3 Performance of the Decomposition with Limited Information

In general, access to news shocks is limited. For this reason, I use the model to propose a series of data moments that can be used to assess the performance of the framework in settings with limited access to news shocks. In particular, I simulate data from the model to study how the method performs in relation to the following two questions: 1) How close is the estimated HEDGE term to its theoretical value?; 2) Are the estimated portable elasticities an upper or lower bound for the theoretical portable elasticities? Throughout the exposition, I consider a researcher who applies the ex-ante decomposition using a single date *zero* innovation in the GE variable.

How close is the estimated HEDGE term to its theoretical value?

I measure the overall discrepancy between the estimated and true HEDGE term as follows.

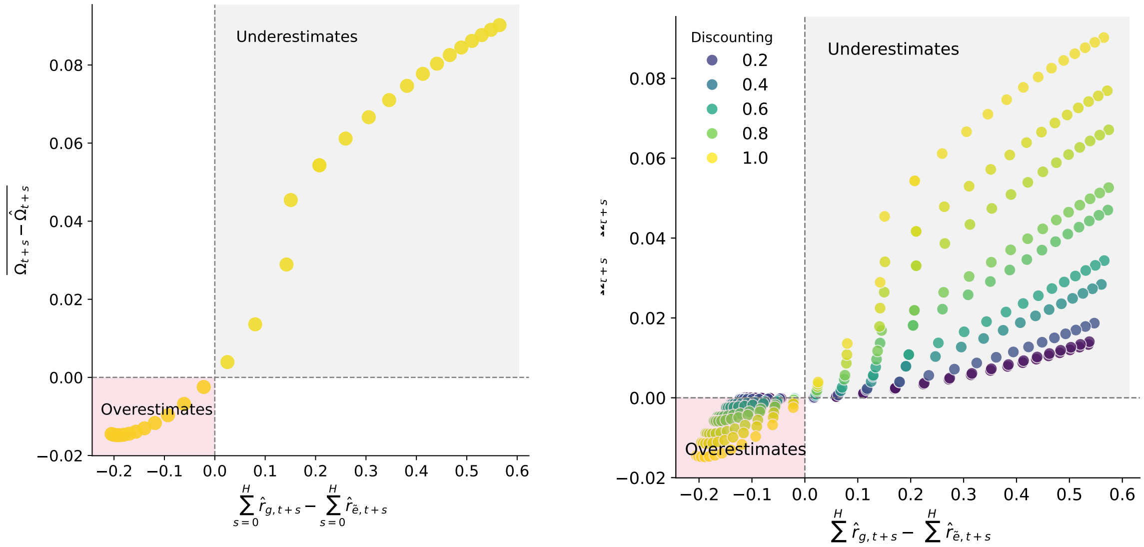
$$D_{\Omega} = H^{-1} \sum_{h=0}^{H} \frac{\Omega_{t+h} - \hat{\Omega}_{t+h}}{\Omega_{t+h}}. \tag{1.99}$$

D_{Ω} measures the average percent deviation between the estimated and true HEDGE term up to horizon H. I express it as a percent to facilitate comparison across calibrations with HEDGE terms of different size.

A sufficient statistic for the overall performance of the decomposition is the difference between the integral of the path for the GE variable in response to the shock of interest and the integral of its counterfactual path using a single date *zero* shock, namely:

$$D_R = \sum_{h=0}^{H} \boldsymbol{R}_{g,h} - \sum_{h=0}^{H} \tilde{\varepsilon}_{r,0} \boldsymbol{R}_{r,h}. \tag{1.100}$$

where $\tilde{\varepsilon}_{r,0}$ is the size of the counterfactual shock that solves the least squares projection in Proposition 1.2. The intuition behind this result is straightforward: the better we enforce the deviation in R from its pre-shock value, the closer the estimation is to the true HEDGE term. Figure 1.10A plots D_R against D_{Ω} for different model calibrations. These calibrations differ only on the assumed persistence for the shock processes. The calibration associated with a smaller D_R yields a better overall performance as measured by D_{Ω}.



(A) Degree of Misalignment - No Discounting. **(B)** Strength of Anticipation Effects.

Figure 1.10. Performance with Limited Information.

Note: The x-axis measures the cumulative differential between the path of R after the shock of interest and the counterfactual implied from a decomposition using only a single date zero shock. The y-axis shows the average (across time) difference between the true model-implied HEDGE term and its estimated value. Positive values on the y-axis imply that on average the estimated HEDGE term is smaller than the true value. Panel (A) plots results for a model without discounting in the Euler equation. Panel (B) plots results for models with different degrees of discounting in the Euler Equation. A higher value (e.g. yellow dots) implies stronger forward-looking dynamics.

Are the estimated portable elasticities an upper or lower bound?

Figure 1.10A shows that the sign of D_R is informative about whether the estimated HEDGE term understates or overstates the true HEDGE term. In particular, when D_R is positive and the deviations of R from the pre-shock values are smaller under the counterfactual than the realized path, then $\hat{\boldsymbol{\Omega}}$, on average, underestimates the true $\boldsymbol{\Omega}$. The reverse is true when D_R is negative. The intuition behind this result is that using a counterfactual path that does not *perturb* the economy enough leads to an underestimation of the HEDGE term. In such a case, the estimated portable elasticity is an upper bound for the true portable elasticity, in absolute terms.

The preceding discussion evaluated performance taking as given the strength of forward-looking dynamics in the economy. Next, I repeat this exercise using simulated data from economies with different degrees of forward-looking forces. I extend the baseline model to allow

Table 1.3. Calibration - Matching D_R in the Model and Data.

Parameter	Description	Value	Target
n	Size of Home Region	.5	Equal sized regions
s_j	Exposure of region $j = H, F$ to ε_g	.8, 1.2	ε_g tilted towards Foreign region
$\bar{\sigma}$	Average IES	1	Nakamura and Steinsson (2014)
σ_H	Home IES	1.2	Home relatively more i sensitive
σ_F	Foreign IES	.8	Foreign relatively less i sensitive
λ_r	Share of *Hand-to-mouth*	.2	Common across regions
ρ_{ε_i}	Persistence of Monetary Policy Shock	.75	Set to match D_R

Note: This table corresponds to the calibration used in Subsection 1.5.4. It details the calibration for parameter values that are either new to the model or that differ from Nakamura and Steinsson (2014). The remaining parameters are detailed in Table A.9.

for a discount term, $\alpha \in (0, 1]$, in the Euler equation (McKay et al., 2017, 2016). The lower α, the more muted the effect of expected future consumption on current consumption decisions - with $\alpha = 1$ in the baseline model. Figure 1.10B shows the simulation results for the same set of calibrations as Figure 1.10A but considering different values for α. First, the two data moments discussed above apply equally well regardless of the strength of forward-looking forces. Second, the plot shows that given a value for D_R, the cost of using limited information decreases as forward-looking forces become less important.

In summary, the previous analysis shows how data moments that are easy to compute can be used to inform the performance of the method, given the modeling assumptions.

1.5.4 Empirical Results Through the Lens of the Model

I take advantage of the mapping between the model and the estimation framework to evaluate the precision of my empirical findings from Section 1.4. Concretely, I calibrate the model to match the empirical value of D_R in order to answer: How well does the decomposition with a single shock perform in this model economy?

To perform the model-based decomposition, I extend the two-region RANK model of Nakamura and Steinsson (2014) to a two-region economy with two types of households: Ricardian and hand-to-mouth (H2M). The environment thus combines the key features of Nakamura and Steinsson (2014) with the two-region TANK model of Herreño and Pedemonte (2025), who analyze heterogeneous regional responses to monetary policy in a setting without fiscal shocks. I assume an equal share of hand-to-mouth households across regions. These households play a key role in disciplining the level of the portable multiplier, which I target to be around one at year two, consistent with the empirical estimates in Section 1.4.

To generate cross-regional heterogeneity in the sensitivity to monetary policy, I introduce region-specific intertemporal elasticities of substitution for Ricardian households. This allows the model to replicate the observed variation in interest rate sensitivity and its correlation with fiscal exposure. In this calibration, the gap between the TWFE and portable cross-sectional multipliers, i.e., the HEDGE term, is driven by the combination of the real rate response to the fiscal shock and the heterogeneity in Ricardians' IES across regions.

In the following, I outline the rest of the calibration details. Most of the parameters are set to their values in Nakamura and Steinsson (2014). These parameters are detailed in Table A.9 in Appendix A.4. Table 1.3 summarizes the parameters that differ from those in Nakamura and Steinsson (2014) or are newly introduced in my extended model. I consider equally sized regions such that $n = .5$. These equally sized regions are heterogeneously exposed to national government spending shocks instead of region-specific government spending shocks. This means that when government spending increases, it increases everywhere but with different intensities across space, in line with the time-series variation used in Section 1.4. I set the exposure shifters to national government spending shocks to $s_H = .8$ and $s_F = 1.2$. Therefore, government spending shocks are biased towards the Foreign region as $s_F > s_H$. These parameters have no direct counterpart in Nakamura and Steinsson (2014) because the authors use region-specific government spending shocks.[37] To generate differential sensitivities to interest rate changes, I allow Ricardian households to have region-specific IES. I set the average IES for Ricardian

[37]Alternatively, one can think of region-specific shocks as setting $s_k = 1$ and $s_{k'} = 0$. In addition, this parameter becomes irrelevant for homogeneous regions, and the only requirement for identification of the cross-sectional multiplier is $s_H \neq s_F$.

households equal to 1 as in Nakamura and Steinsson (2014) and the region-specific IES to $\sigma_H = 1.2$ and $\sigma_F = .8$. The Home region is relatively more sensitive to changes in the interest rate. This calibration reproduces the correlation between the sensitivities to fiscal and monetary policy shocks found in Section 1.4. This exercise targets the sign of the correlation between fiscal and monetary sensitivity rather than the dynamic path of cross-sectional interest rate responses, a task that I leave for future work. Lastly, the share of hand-to-mouth households is set at $\lambda_r = .2$ for both regions to target the level of the portable multiplier close to unity on the two-year horizon.

Crucial for this exercise are the parameters that govern the response of the real rate to government spending and monetary policy shocks. To minimize my departure from the calibration in Nakamura and Steinsson (2014), I match D_R by choosing the persistence for the monetary policy shock, leaving the process for government spending unchanged. This requires the monetary policy shock to follow an AR(1) with persistence $\rho_{\varepsilon_i} = .75$. Formally,

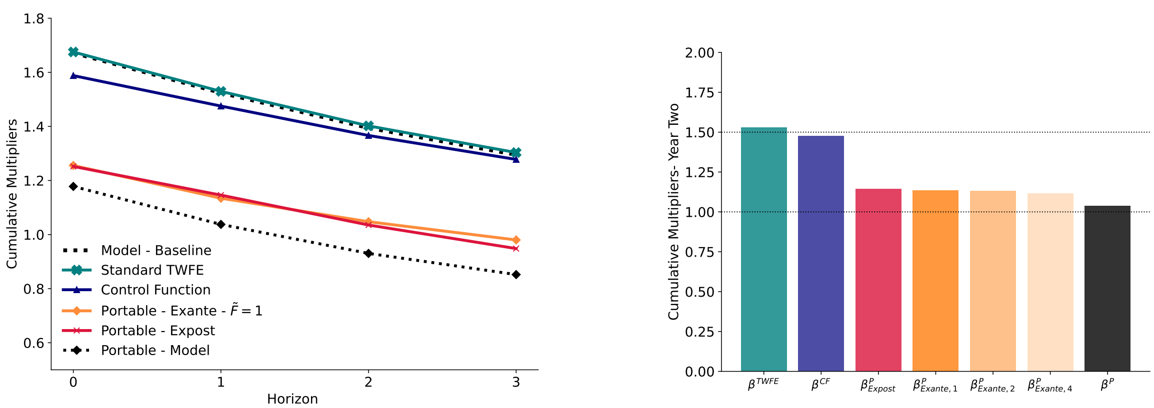
$$i_t = \rho_i i_{t-1} + (1-\rho_i)[i_{ss} + \phi_\pi \Pi_t^{agg} + \phi_y \hat{Y}_t^{agg}] + u_{it}, \qquad u_{it} = \rho_{\varepsilon_i} u_{it-1} + \varepsilon_{it}. \tag{1.101}$$

The nominal interest rate follows an Taylor rule subject to an AR(1) shock, and national government spending follows an exogenous AR(1) process, as in Nakamura and Steinsson (2014).

Next, I simulate data from the model and apply the decomposition framework assuming that only a contemporaneous interest rate shock is available. Figure 1.11 displays the results. The left panel plots the cumulative fiscal multipliers across horizons, comparing the true model-implied portable multiplier (black dotted line) with several alternative estimators. Both the TWFE estimator and the control function approach exhibit upward bias relative to the true portable multiplier, driven by the positive HEDGE term. In contrast, estimates obtained by applying either the ex post decomposition or the ex ante decomposition with a single shock correct for most of the omitted variable bias, bringing estimated multipliers much closer to the true portable effect.

Importantly, both versions of the decomposition perform well even in a setting with strong forward-looking dynamics (McKay et al., 2016; Del Negro et al., 2023). This is consistent with the findings of the previous subsection, which showed that a small D_R - that is, a small discrepancy between the realized and counterfactual path of R -implies a limited precision loss under restricted information.

Figure 1.11B zooms in on the year-two multiplier and quantifies the extent to which each method corrects for the HEDGE term. While the TWFE and control function estimators overstate the true portable multiplier by approximately 0.5 points, the decomposition-based estimators align much more closely with the model-implied benchmark. In particular, allowing for additional news shocks to the monetary policy rule ($F = 2$ or $F = 4$) yields only modest improvements relative to the single-shock decomposition. This highlights that when the counterfactual path of R can be closely approximated, even with limited instruments, the decomposition method delivers accurate estimates of the portable elasticities.



(A) Estimated Multipliers in the Model.

(B) Zoom in - 2-year Cumulative Multiplier.

Figure 1.11. Decomposition Results using Model Simulated Data.

Note: Panel (A) shows the estimates of the cross-sectional fiscal multipliers for different estimation strategies, together with the true model implied values. Panel (B) zooms in at the 2-year horizon multipliers shown in the left panel. Results are based on 100 repetitions.

1.6 Conclusion

This paper addresses the identification challenge that arises in cross-sectional empirical strategies when units' exposure to multiple aggregate variables is driven by common exposure shifters. In such scenarios, the typical empirical design that relies on cross-sectional variation in exposure combined with time fixed effects fails to separately identify portable or partial equilibrium effects from general equilibrium driven responses. This challenges the *portability* of these estimated elasticities, limiting their external validity and their usefulness for policy analysis.

To overcome this problem, I introduce a decomposition framework that explicitly separates portable effects from general equilibrium driven effects by jointly exploiting cross-sectional and time-series variation. The method clarifies the conditions under which portable statistics can be consistently estimated and implemented under varying informational assumptions ranging from static to forward-looking dynamic environments.

Applying this framework to US data, I estimate the cross-sectional fiscal multiplier and demonstrate that accounting for general equilibrium effects operating through monetary responses significantly reduces the estimated cross-sectional multiplier. This finding challenges the conventional assumption in this literature that monetary policy responses are fully absorbed by time-fixed effects, highlighting the importance of explicitly accounting for monetary-fiscal interactions when estimating cross-sectional multipliers.

More broadly, the framework can be applied beyond fiscal policy to settings where cross-sectional estimates may be shaped by general-equilibrium feedback—such as monetary, trade, or credit-supply shocks. By clarifying how to recover portable elasticities in HEDGE economies, it provides a foundation for reinterpreting and extending existing work that combines micro data with aggregate shocks. Future research can use this approach to obtain clean empirical counterparts of the objects required for model calibration, policy evaluation, and counterfactual analysis.

Acknowledgements

Chapter 1, in full, is based on the paper "Cross-sectional Identification with Heterogeneous Exposure to General-equilibrium Effects" currently being prepared for submission. The dissertation author is the sole author of this paper.

Chapter 2

The Financial Channel of Tax Amnesty Policies

2.1 Introduction

The size and depth of a country's financial system are central to fostering innovation, facilitating international trade, and promoting economic development. However, in many economies, a substantial share of domestic savings is concealed offshore or held informally outside the banking system to evade taxes (Alstadsæter et al. (2018)). This practice undermines both public revenue collection and the development of domestic financial markets.[1] If even a small portion of this wealth were repatriated, it could generate a sizable inflow of foreign-currency deposits into domestic banks. Yet little is known about the policies that can trigger such repatriation and about what would be the real effects of such a sudden increase in banks foreign-currency liquidity.

One policy that could play this role is a tax amnesty—a temporary program allowing taxpayers to declare previously untaxed or hidden assets with reduced penalties. Since 2005, more than 40 countries have implemented a tax amnesty policy. Notable examples - each revealing assets exceeding \$50 billion - include those launched in Brazil (2016), Italy (2009), Argentina (2016) and Indonesia (2016). Recent literature has primarily focused on the fiscal channel of

[1] Alstadsæter et al. (2018) documents that approximately 10% of global GDP is held in tax havens. This estimate is likely conservative, as it does not account for cash holdings concealed outside the formal banking system, such as cash stored at home or in security boxes.

tax amnesties, examining how asset disclosures affect governments' ability to raise revenue. However, one potentially important aspect remains overlooked: the capacity of tax amnesty policies to expand the economy through the financial system. By channeling hidden assets into domestic banks, these policies can increase bank liquidity and the supply of bank credit and, in turn, promote investment and growth of firms in the private sector. We refer to this mechanism as the *financial channel* of tax amnesty policies. Despite its potential relevance, particularly in economies with underdeveloped financial systems, there is currently no evidence documenting this channel.

This paper studies the effects of large scale asset repatriations in the context of Argentina's 2016 Tax Amnesty, one of the largest asset disclosure programs in history. The program generated an influx of savings into domestic banks equivalent to 1.4% of GDP, the vast majority of which were denominated in US dollars. We exploit the heterogeneous exposure of domestic banks to these inflows, combined with variation in firm exposure to affected banks, to identify the response of banks and firms to the amnesty-induced financial shock. We find that more exposed banks significantly increased their credit supply, leading to improvements in firms' financial and real outcomes - particularly in overall bank borrowing, access to dollar credit, imports, exports, and employment. These results suggest that the gains from tax amnesties through the financial system can be substantial, especially in shallow credit markets. Moreover, because the shock shifted the relative supply of dollar- versus peso-denominated credit, we use this episode to study firms' currency composition decisions and how these affect their input mix in the production function. We find evidence of substitution between debt in different currencies and significant heterogeneity between firms. To quantify the macroeconomic importance, we develop a small open economy model with heterogeneous firms that finance working capital through a monopolistically competitive banking sector. In the baseline calibration, the model implies that the tax amnesty increased GDP by 0.9%.

The Argentinian experience provides an ideal setting for studying the effects of large-scale capital inflows and the *financial channel* of tax amnesty policies for several reasons. First, the program was remarkably successful, leading to the disclosure of assets equivalent to 21% of GDP[2]. The program allowed citizens to declare previously hidden assets in exchange for a relatively small penalty. Crucially, it offered incentives to repatriate financial assets from foreign bank accounts to domestic banks. As a result, more than 20% of all disclosed bank holdings were deposited in Argentine banks, an amount equivalent to 1.4% of GDP[3]. Second, the vast majority of repatriated savings were denominated in US dollars. In this respect, the tax amnesty resembled an unusual episode of reverse capital flight, offering a unique opportunity to study the impact of foreign currency inflows on the domestic financial system, including changes in bank and firm financial behavior as well as firms' participation in international markets. Third, Argentina is a middle-income economy with a relatively underdeveloped financial system and limited access to credit markets. At the time of the amnesty, domestic bank credit to the private sector was only 13% of GDP, well below the Latin American average of 40%, making the economy particularly suited to benefit from an expansion in credit supply.

Studying the financial channel of tax amnesty policies has been challenging because it is difficult to combine policy variation with comprehensive data. We overcome these challenges by combining bank-to-firm lending data from the Credit Registry of the Central Bank of Argentina with firm-level employment data from the tax authority and trade data from customs records. Our dataset covers the universe of firm-to-bank credit transactions, the currency composition of firm-level credit, bank balance sheets, and firm-level outcomes. A key feature of our data is that it identifies the precise amount of dollar deposits each bank received as a result of the tax amnesty. This level of granularity is rarely available to researchers and is specific to Argentina, where regulatory mandates required amnesty-related deposits to be made into designated special bank accounts.

Our empirical analysis is structured around four steps. First, we assess the effectiveness

[2]This figure includes both real (e.g. residential properties, businesses, vehicles, art, etc.) and financial (e.g. cash, bank accounts abroad, equity holdings, etc.) assets.

[3]Cash and bank account holdings accounted for 30% of total disclosed assets, making them the second most significant asset category. The first asset category in terms of disclosed values was portfolio investments.

of the tax amnesty policy in inducing citizens to deposit concealed cash holdings in domestic banks. Second, we analyze how banks responded to the inflow of tax amnesty dollars. Third, we investigate firm-level borrowing responses as well as the effect of the shock on firms' currency choices. Finally, we examine the real effects of the tax amnesty policy, focusing on how it affected firms' performance in labor and international markets.

We start by documenting the significant and unprecedented inflow of dollars into the Argentine banking system due to the tax amnesty. In a two-month window, dollar deposits in domestic banks increased by 50%, while peso deposits remained largely unaffected by the policy. Furthermore, we show that there was considerable variation in how tax amnesty dollars flowed into banks. Banks with a higher proportion of dollar deposits in their liabilities before the amnesty experienced stronger inflows of dollar funds. Although median bank deposits increased by 2% of their total private sector deposits, some banks experienced increases over 13%, and others did not receive additional deposits.

We study the response of bank lending by leveraging this variation in bank exposure to the tax amnesty. We measure bank-level exposure as the ratio of dollar deposits received during the tax amnesty to the bank's total deposits. Importantly, we show that, across exposure levels, banks were balanced in terms of both observable characteristics and pre-amnesty lending trends. To identify the effect of the tax amnesty on bank lending, we employ the within-firm approach popularized by Khwaja and Mian (2008) that leverages variation in exposure across different banks lending to the same firm. By only exploiting within-firm variation, our estimation strategy effectively controls for any potential correlation between firm-specific credit demand shocks and the distribution of bank exposure. We find that bank exposure resulted in a relative increase in lending. On average, banks in the 75th percentile of exposure granted 11% more credit relative to those in the 25th percentile. This effect is more pronounced for credit granted to exporters, which are the firms that typically rely more on dollar-denominated debt.

Next, we study the impact of the tax amnesty on the financial and real performance of firms. For identification, we leverage the fact that not all firms were equally exposed to the increase in credit supply since they were borrowing from different banks before the tax amnesty. We construct firm-level exposure as a weighted average of bank exposure, where the weights correspond to the share of debt with each bank prior to the tax amnesty. The intuition is that there is some rigidity in firm-bank relationships, so a firm is more exposed to the shock if it is already connected to those banks that receive more dollars. Our identification assumption is that, in the absence of the tax amnesty, firms connected to high-exposure banks would have followed a similar trend to firms not connected to them. We provide evidence supporting this assumption by showing that the pre-trends between more exposed and less exposed firms are not significantly different prior to the tax amnesty. In addition, we restrict our comparisons to firms with heterogeneous exposure levels within narrowly defined industries and with the same trade status. This approach helps to control for potential confounding effects from any post-amnesty shock that might differentially affect exposed and non-exposed firms due to their preexisting characteristics.

We first test whether the positive effects observed in the *within-firm* analysis extend to the total borrowing of the firms, as it is possible that firms were mainly substituting existing debt across banks. Our findings provide evidence against such substitution. More exposed firms experienced greater credit growth compared to less exposed ones. Our estimates are of similar magnitude to those from the within-firm analysis, which we interpret as evidence of a strong firm borrowing channel rather than a reallocation of existing borrowing to more exposed banks. Specifically, we estimate that an average firm in the 75th percentile of exposure experienced a 9.7% increase in credit relative to a firm in the 25th percentile of exposure. Consistent with our within-firm results, we also find stronger effects among differentially exposed exporters, with the impact of the tax amnesty being more than twice as large as that observed for the average firm.

Next, we evaluate the performance of the dollar credit market in response to the tax amnesty shock, which improved the capacity of banks to issue dollar-denominated credit. In emerging economies, the availability of foreign currency credit, particularly in dollars, matters for firm performance. Fluctuations in dollar credit availability have been shown to lead to reduced firm investment (Kalemli-Ozcan et al. (2016),Elias (2021)), reduced export dynamism (Paravisini et al. (2015b)), and weaker employment growth (Hardy (2023)). Thus, it is important to assess how tax amnesties, when they channel foreign-currency savings toward domestic banks, contribute to smoothing and deepening access to dollar credit. Given that few firms held dollar debt at the beginning of the sample, our analysis focuses on the extensive margin, examining how changes in foreign currency funding impact firms' participation in this credit market. Our results indicate that firms that were more exposed to the tax amnesty were significantly more likely to borrow in dollars. Specifically, firms in the 75th percentile of exposure had a 3 percentage point higher probability of borrowing in dollars compared to those in the 25th percentile. This represents a 10% increase relative to the unconditional probability of holding dollar debt at the start of the sample.

We then examine how the tax amnesty influenced the currency composition of firms' borrowing, shifting our focus from credit growth outcomes to the currency choices made by borrowing firms. Understanding the ease with which firms can substitute between credit in different currencies is crucial to assessing the real and aggregate consequences of fluctuations in dollar credit availability, particularly in emerging economies. We provide a causal estimate of the elasticity of substitution between peso and dollar debt by leveraging the currency-specific nature of the shock. We argue that the amnesty created a heterogeneous firm-level shifts in the relative supply of dollar-denominated loans, offering a promising quasi-experimental setting to explore this substitution elasticity. We discuss how the institutional regulations of the financial system are essential for this supply shifter to work as proposed and conduct tests that show no differential movements in peso lending rates between more and less exposed banks.

We estimate this elasticity of substitution using firm-level exposure to the tax amnesty as a shifter in the relative cost of debt in different currencies. We find that peso- and dollar-denominated credit are at least weak substitutes, with differential effects between exporters and non-exporters. Although understanding the drivers of this heterogeneity is beyond the scope of our paper, our results suggest a role for differences in productive processes and tasks in shaping firms' debt portfolio decisions. We believe that investigating what shapes this heterogeneity is a promising research avenue, as it can provide insights into how these differences might affect aggregate outcomes when dollar credit fluctuates.

Lastly, we examine how easing credit access affects firm behavior across several dimensions. Firms with greater exposure to the amnesty shock experienced a relative increase in total imports, particularly of intermediates and capital goods, highlighting a direct impact on investment. Exports also rose significantly for more exposed firms relative to less exposed ones, with those in the lowest export quartile increasing exports by 4% for each percentage point of exposure. Employment increased by an estimated 2% per percentage point of exposure, indicating that credit access supports operational expansion and job creation. However, we find no evidence of a relative change in average wages, suggesting that increased labor demand was met by a relatively elastic firm-level labor supply. Taken together, these results underscore the role of credit access in enhancing firm growth, resilience, and market integration.

To interpret the empirical findings and assess their aggregate implications, we develop a small open economy general equilibrium model with heterogeneous banks and multi-currency firm financing. The model builds on the framework of Herreño (2023), extending it to an open economy environment with multiple financing currencies. As in Herreño (2023), imperfectly competitive banks set lending rates and firms finance working capital needs through bank credit. We enrich this framework by allowing banks to raise funding in multiple currencies and firms to choose both the bank and the currency of borrowing. The model features banks that set currency-specific lending rates, where the marginal cost of dollar wholesale funding is decreasing in the bank's dollar deposit base. On the firm side, producers use labor and imported intermediate inputs, and must finance in advance each input's cost through working capital loans. The currency composition of financing needs is allowed to differ across inputs. Specifically, guided by the data, imported inputs are assumed to be more dollar-intensive than wage payments, reflecting that international transactions are invoiced in dollars and that domestic banking regulations channel dollar credit toward trade-related activities. This generates a direct link between currency-specific deposit shocks and firms' relative input costs: an inflow of dollar deposits compresses dollar lending rates, disproportionately reducing the effective cost of imported inputs relative to labor, and thus reshaping the optimal input mix.

The model delivers a mapping between our empirical reduced-form coefficients and structural objects, which we exploit for identification and calibration. In particular, we show how the relative reduced form responses of the dollar share, imports and labor help discipline the elasticity of substitution between peso and dollar debt. In our preferred calibration, we estimate aggregate effects of the tax amnesty shock equivalent to 0.9 percent of GDP. This implies a GDP-to-deposits aggregate elasticity of around .13.

In summary, these results suggest that tax amnesty policies can expand a country's financial sector, increase access to credit, and foster firm growth in both domestic and international markets by incentivizing the flow of private savings into domestic banks. Policymakers gain two key insights: first, tax amnesty policies have benefits beyond raising government revenue; second, these policies should be designed to not only encourage asset disclosure, but also actively promote the deposit of cash holdings into domestic banks. This is important since many tax amnesty laws stop at disclosure incentives and fail to create channels that direct funds into the domestic financial system.

More broadly, our results highlight that imperfect substitution between dollar and peso debt, combined with the differential specialization of currencies in financing different production inputs, implies that the real effects of capital flow episodes depend critically on the currency denomination of the shock. When liquidity expansions are dollar-specific the transmission to production decisions is asymmetric: if international trade financing is more dollar-intensive than wage-bill financing, a dollar liquidity shock disproportionately reduces the effective cost of imported inputs relative to labor, tilting firms' optimal input mix toward imported intermediates.

Related Literature

This paper contributes to four strands of the literature. First, the paper relates to recent literature that studies the consequences of tax amnesties. Previous studies in this literature have primarily focused on the immediate fiscal impacts of these policies on disclosure of hidden assets and its effect on government tax revenues and public spending (Londoño-Vélez and Tortarolo, 2022; Langenmayr, 2017; Lejour et al., 2022; Gil et al., 2024). For instance, Londoño-Vélez and Tortarolo (2022) investigate the fiscal impact of the same policy in Argentina and find that the tax amnesty led to improved tax compliance, increased government revenue, and expanded social transfers. Gil et al. (2024) study the importance of deterrence laws and beliefs about future amnesties on incentivizing asset disclosure. Unlike these previous studies, which focus on hidden asset declarations and their fiscal effects, our paper is the first to examine the indirect consequences of the policy through the financial sector and spillovers to firms. We show that tax amnesties can encourage the deposit of hidden assets into domestic banks, and that banks more exposed to these inflows expand their credit supply. Consequently, firms connected to these banks experience improved performance due to enhanced access to credit. Our findings suggest that tax amnesties have the potential to influence economic growth through their effects on banks and firms in the private sector, extending beyond their direct impacts on government tax revenues and spending.

Second, this paper contributes to the literature in macroeconomics that identifies the economic effects of credit supply shocks by isolating the bank lending channel. Papers in this strand include Kashyap and Stein (2000); Amiti and Weinstein (2011); Khwaja and Mian (2008); Schnabl (2012); Federico et al. (2023); Mora (2013); Chodorow-Reich (2014); Herreño (2023); Paravisini (2008); Villacorta et al. (2023); Paravisini et al. (2015b); Alfaro et al. (2021); Amiti and Weinstein (2018); Jiménez et al. (2020); Aristizábal-Ramırez and Posso (2024), among others. We focus on a relatively underexplored dimension of the bank lending channel: the currency denomination of credit supply shocks. By exploiting the regulatory framework of Argentina's tax amnesty, we identify a shock that expanded dollar liquidity across banks heterogeneously, without a corresponding shift in peso funding conditions. This allows us to study not only how banks adjust their lending in response to a foreign-currency deposit shock, but also how the currency composition of credit supply propagates to firms' real decisions. Crucially, we show that the currency of the shock matters for transmission: because imported inputs are more dollar-intensive, a dollar liquidity shock disproportionately reduces the effective cost of importing relative to hiring, reshaping firms' optimal input mix.

We also contribute to this literature by extending the framework of Herreño (2023) to an open economy with multi-currency financial frictions. The model links the currency composition of bank funding to firms' input demand and financing decisions, allowing us to study how the denomination of credit supply shocks shapes real economic activity.

Third, this paper relates to the literature on financial development, credit constraints, and firms' performance (Levchenko et al. (2010); Chor and Manova (2012); Caggese and Cuñat (2013); Manova (2012, 2008); Paravisini et al. (2015b,a); Federico et al. (2023); Kohn et al. (2016); Leibovici (2021); Kohn et al. (2023)). While most of these studies focus on the role of access to credit in facilitating exporting, our paper extends the analysis to the causal impact on importing. In this context, Kohn et al. (2023) empirically document that financially underdeveloped economies exhibit a slower aggregate response following trade liberalization, attributing this to credit constraints affecting importing possibilities. Additionally, Muûls (2015) documents the relationship between importing and access to credit. We contribute to this literature in two dimensions. First, we combine exposure of banks to the tax amnesty with firm-to-bank data to empirically estimate the causal effect of credit availability, especially in dollars, on firms' importing activities. Second, most of these papers do not have access to data on the currency denomination of the credit. We are able to study how access to dollar credit deferentially affect firms' participation in international markets.

Fourth, this paper contributes to the literature on flight capital and capital reversals. In this field, commercial banks have been identified as significant channels of contagion during periods of international capital reversals, as they often reduce their credit supply, negatively impacting firms' borrowing and output. Since sudden inflows of capital are hard to identify, most studies focus on crises and sudden capital outflows. In contrast, our paper adds to this literature by documenting the effects of a sudden inflow of capital into banks and examining how this influx influences the broader economy.

Finally, the paper is also related to the literature exploring the drivers of firms' foreign currency debt choices in emerging economies (such as Calvo (2002); Galindo et al. (2003); Allayannis et al. (2003); Kamil (2012); Brown et al. (2011); Degryse et al. (2012); Brown et al. (2014); Hardy (2018)). The richness of our data and the regulatory environment of the tax amnesty provides an ideal scenario to show novel patterns about the relationship between dollars and pesos debt. In addition, the dollar-specific shock to banks allow us to identify the firm borrowing elasticity between pesos and dollar debt.

The remainder of the paper is organized as follows. Section 2.2 provides background on tax amnesty programs and details the institutional context of the Argentine case. Section 2.3 describes the datasets used in the empirical analysis. Section 2.4 outlines the identification strategy and presents the main results. Finally, Section 2.6 offers concluding remarks.

2.2 Tax Amnesties and Argentina's Experience

Tax Amnesties

A tax amnesty, as defined by Le Borgne and Baer (2008), is a limited time offer by the government that allows taxpayers to settle their tax liabilities for a defined amount, often including a reduction or forgiveness of interest and penalties. This settlement pertains to previous tax periods and grants taxpayers freedom from legal prosecution related to the disclosed tax liabilities. Le Borgne and Baer (2008) highlight three goals of these programs: (1) raise government revenue quickly, (2) increase future tax compliance, and (3) incentivize asset repatriation of flight capital for reasons that go beyond immediate revenue and tax compliance motives.

The existing literature has primarily addressed the first two objectives, largely overlooking the indirect effects through asset repatriation. This paper focuses on asset repatriations. We adopt a broad definition of "asset repatriation," which includes inflows into domestic banks from: (a) cash previously held abroad, and (b) cash held domestically but informally outside the banking system (for example, “under the mattress” or in private vaults).[4]

Tax amnesty programs are increasingly used by governments as policy tools, a trend expected to grow due to enhanced measures against cross-border tax evasion, making these programs more likely to succeed.[5] In the past 20 years, over 30 countries have launched tax amnesty programs with varying levels of success in terms of asset disclosure. Notable examples of successful tax amnesty programs include Brazil (2016) with $50 billion disclosed, Chile (2015) with $18.7 billions, Italy (2009) with $102 billion, Indonesia (2016) with $346 billion, Argentina (2016) with $116 billion, and Pakistan with $21 billion. Other countries that implemented national programs are United Kingdom, Russia, Spain, India, among others. Regarding asset repatriation, some of these programs included incentives for repatriation, while others did not. Unfortunately, data on the exact amount of repatriated cash holdings for most of these programs is typically not available.

Argentina’s Tax Amnesty

In July 2016, Argentina enacted a tax amnesty law allowing citizens who had previously concealed assets from the tax authority to disclose them by paying a modest penalty for past tax evasion. This policy was part of a larger government initiative aimed at achieving budget balance. Specifically, the tax amnesty played an important role in immediately increasing tax revenue through the collection of one-time penalties and by broadening the future tax base.

Argentina’s 2016 tax amnesty is one of the largest of its kind worldwide. Disclosed assets amounted to approximately 19% of GDP, effectively doubling the country’s wealth tax base. Londoño-Vélez et al. (2022) provide a detailed analysis of the fiscal aspects of this episode. In the following, we highlight the aspects of the amnesty that are particularly relevant for the present

[4]Although informally held domestic cash is already within the country, its effects resemble those of repatriated funds, as both increase banks’ dollar deposit base.

[5]Examples of such measures include the US Foreign Account Tax Compliance Act (FATCA) and the global Common Reporting Standard (CRS) for the automatic exchange of financial account data.

study.

Under the policy, citizens could disclose any type of asset subject to wealth taxes, regardless of its geographical location (whether domestic or offshore). Eligible assets included stocks, portfolio investments, residential properties, cars, foreign currency checking accounts, and cash holdings, among others. A significant portion of the disclosed assets were held abroad, notably foreign currency checking accounts, real-estate, and financial investments.

In this study, we focus on the disclosure of two types of financial assets: (i) foreign-currency savings held in offshore bank accounts and (ii) foreign-currency cash holdings stored informally outside the domestic financial system (e.g., cash stored at home or in safety deposit boxes). We bundle them together as "repatriated cash holdings" because both increase the dollar deposits in the domestic financial system.

Specifically, citizens who disclose assets held in foreign bank accounts were given two options: keep these assets abroad or repatriate them to a domestic bank. Repatriation was incentivized by a lower penalty rate, encouraging inflows of these funds into Argentina's banking system. In contrast, citizens who disclosed foreign currency cash holdings were required to deposit them in domestic banks, as there was no alternative option available. Both types of assets had a defined time window, August to December 2016, for their repatriation or formal deposit into the domestic banking system. Thus, from the perspective of domestic banks, these two distinct asset classes represented an economically equivalent inflow of foreign currency deposits during the amnesty period.

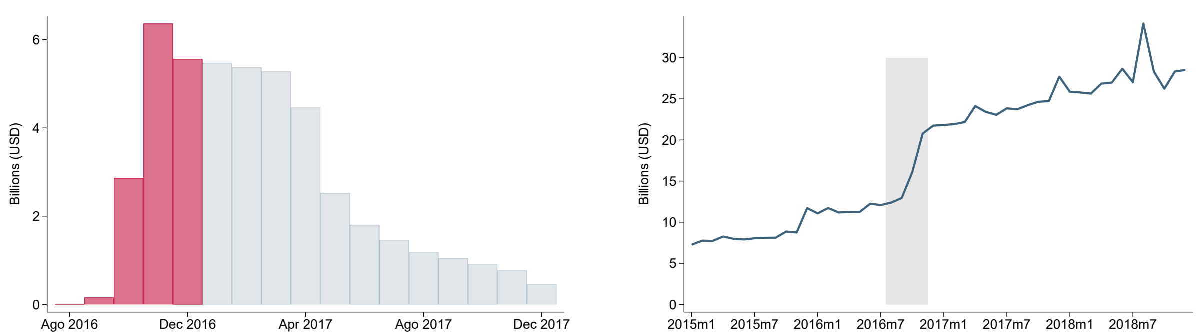
Citizens who opted to declare cash or repatriate foreign currency had to open a special foreign currency bank account at a bank of their choice to deposit the money. This is convenient because it allows us to observe the exact amount of dollars that entered the bank system because of the tax amnesty. We will refer to the specially created bank accounts as *Tax Amnesty Accounts*. In addition, they were required to park declared funds in those same accounts for six months[6].

[6]There were a number of exceptions to the parking requirement. For example, the requirement could be lifted if funds were applied to purchase government debt securities

The amount of foreign currency deposits (mainly US dollars) that flowed into domestic banks was equivalent to 1.4% of GDP and implied almost a doubling of the stock of private dollar deposits in domestic banks. This occurred within a very short window of around three months as most funds were effectively repatriated between October 2016 and December 2016. This inflow of dollars into domestic banks is what we refer to as the tax amnesty shock throughout the paper.[7]

Figure 2.1A shows the aggregate stock of dollar deposits held in *Tax Amnesty Accounts*. The inflow of tax amnesty funds was concentrated in the last quarter of 2016 even though the tax amnesty window opened in August 2016. After the 6-month parking period, depositors could transfer their holdings to their standard dollar checking accounts, which explains the gradual decline in deposits in Figure 2.1A.

Figure 2.1B contextualizes the relevance of the tax amnesty shock. It plots the evolution of total private sector dollar deposits. The tax amnesty resulted in a significant increase in dollar-denominated deposits at domestic banks, which doubled. Tax amnesty funds explain more than 80% of the increase in dollar deposits during the last semester of 2016.



(A) Deposits in *Special Accounts*

(B) Dollar Deposits

Figure 2.1. Tracking Tax Amnesty Dollars

Note: Data are in billions of US dollars. Panel (a) shows the evolution of dollar deposits in the specially created bank accounts. The red bars correspond to the months of the tax amnesty window. Each bar corresponds to the aggregate stock of dollar deposits at the end of each month. After the 6-month parking period, funds gradually flow out of the special accounts into traditional dollar checking accounts. Panel (b) shows the evolution of dollar deposits held by the private sector. The shaded region corresponds to the tax amnesty window.

[7]Note that this excludes from the analysis any foreign currency holdings at bank accounts abroad that citizens did not choose to deposit in Argentinian banks.

2.3 Data

We combine four datasets: i) bank-level balance sheet data obtained from the Central Bank of Argentina (BCRA); ii) firm-level data on exports and imports from the Argentinian Customs Office; iii) firm-to-bank data on domestic credit compiled by the Central Bank of Argentina (BCRA); and iv) administrative records on firm-level employment and average wages from Argentina's tax authority (AFIP). All data sets are collected monthly and cover years from 2014 to 2018. We employ bank-level balance sheets to recover information on banks' assets, liabilities, core capital, and profits. These data cover a total of 80 banks operating in Argentina, both publicly and privately owned. We complement balance sheets with bank-level performance indicators (e.g., leverage ratios, return on assets, and average interest rates), which are published by the Central Bank. Importantly, the balance sheet records the amount of dollar deposits that each bank received during the tax amnesty. Banks were required to inform the outstanding stock of deposits in checking accounts that were specially created for tax amnesty.[8]We restrict the analysis to banking institutions that had a participation in total deposits of at least 0.05% during our sample period. This leaves us with 28 domestic banks that take up more than 97% of aggregate deposits.

We use Credit Registry data to obtain firm-level credit indicators. We employ two data sets within the Credit Registry. Our primary source of information contains detailed firm-bank level data on monthly total outstanding loans. This dataset covers all private non-financial firms that have debt with domestic banks between 2014 and 2018. We complement this information with a second dataset that started being recorded in August 2015. This additional data set contains disaggregated information on the currency composition of firm-level debt. For each firm, we observe the level of pesos and dollar debt over time. We focus on firms that report debt to banks at some point during the period considered. In addition, we drop firms for which total credit

[8]Section 2.4 provides more details on how we use this information to construct the exposure of banks to the tax amnesty shock.

never exceeds 50 US dollars and winsorize the top and bottom 1% of observations to limit the influence of outliers. After this process, we are left with data on 77,777 firms.

Lastly, we gather information from several sources to obtain a comprehensive picture of the consequences of the tax amnesty on firms' real performance. We use comprehensive customs data covering the universe of firm-level export and import transactions. This data set covers the period 2014 to 2018. It provides monthly information on the value (in US dollars) of exports and imports for each firm, categorized by country (origin/destination) and product at the 6-digit level. We link this data set to fiscal files generated by the Argentina Tax Authority (AFIP) using unique firm identifiers to enrich our analysis. This allows us to obtain additional information on formal employment, average wages, and firms' main sector of activity. Consequently, we are able to construct firm-level measures of employment, exports, imports, and payroll for the years 2014 to 2018.

2.4 Empirical Analysis

We empirically examine the impact of the tax amnesty in Argentina through the financial channel. Specifically, we analyze to what extent the resulting capital inflow influenced banks' credit supply and, in turn, how the expanded credit access affected output of firms.

We first construct bank- and firm-level exposure measures to the tax amnesty. The empirical analysis then proceeds in two steps. First, we examine the bank lending channel, assessing whether more exposed banks expand their credit supply by more. Second, we analyze the impact of the increase in credit supply on firm-level outcomes, including borrowing, trade (imports and exports), and employment.

2.4.1 Bank and Firm Exposure

Bank-level Exposure

The balance sheet of domestic banks records the stock of deposits held in the special accounts created during the tax amnesty. For each bank b, we compute the average deposits reported in these special accounts between October 2016 and December 2016. We refer to this object as the *tax amnesty funds* received by bank b. Next, we measure the exposure of bank b to the tax amnesty as the ratio between its *tax amnesty funds* and its total private sector deposits in the quarter preceding the tax amnesty. Formally,

$$S_b = \frac{Tax\ Amnesty\ Funds_b}{D_{b,0}^{Total}} \times 100, \tag{2.1}$$

where $D_{b,0}^{Total}$ are the average total deposits in 2016Q2. S_b measures the importance of the inflow of dollars due to the tax amnesty relative to the existing liquidity of each bank. Banks experienced varying levels of exposure to the tax amnesty funds, with an average exposure of 10% and a standard deviation of 10 percentage points. For the most exposed banks, the liquidity shock represented nearly 20% of their preexisting deposit liabilities. The difference in exposure between a bank in the 75th percentile of exposure and one in the 25th percentile of exposure is 10 percentage points.

The differences in inflows at the bank level were partly shaped by the existing banking relationships of households. Three pieces of evidence support this. First, the variation in bank exposure is positively correlated with the share of dollar deposits a bank held before the amnesty (Appendix Figure B.9). Second, at the municipality level, inflows are concentrated in areas with higher pre-existing dollar deposits, indicating households deposited funds in banks where they had prior relationships. Third, most individuals transferred funds from special amnesty accounts to other accounts within the same bank after the holding period ended, suggesting that they already had accounts there (Appendix Figure B.10).

Although our identification strategy does not rely on balance in pre-amnesty characteristics, examining the relationship between these characteristics and our exposure measure is informative. Figure 2.2 presents the coefficients from individually regressing a comprehensive set of baseline bank characteristics on tax amnesty exposure. With the exception of public ownership, for which we explicitly control in all specifications, we find no systematic differences between banks with higher and lower exposure.

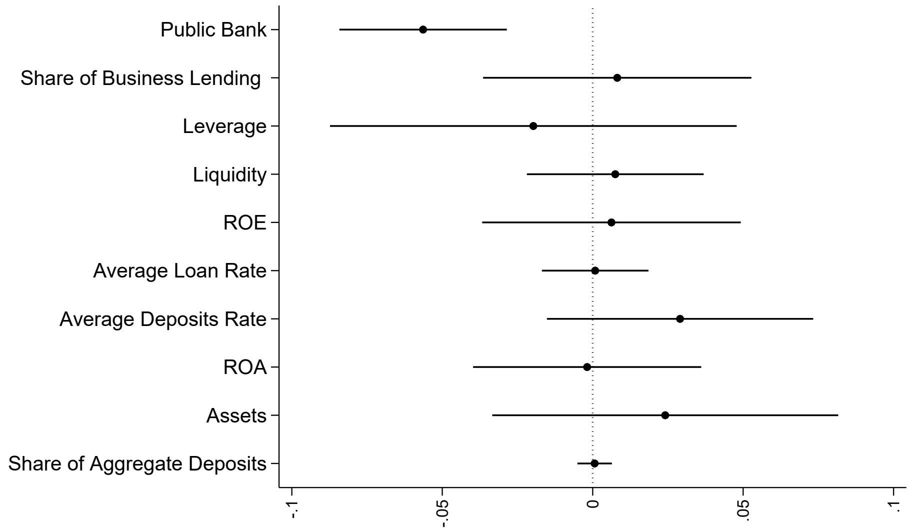
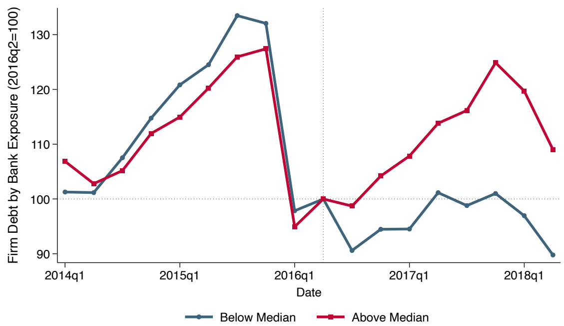


Figure 2.2. Banks Characteristics - Balance Test

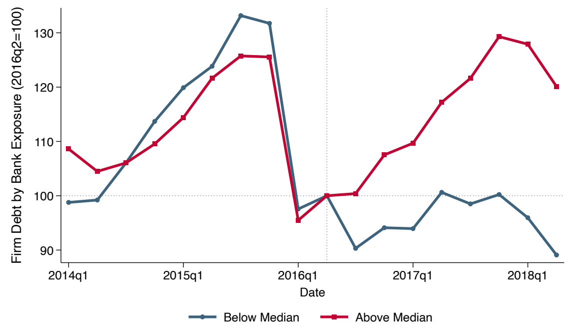
Note: We report 95% confidence bands. Bank characteristics are averages of the year before the tax amnesty. All variables were standardized with the exception of the participation of each bank in aggregate deposits (our *proxy* for bank size). Regressions are weighted by the share of each bank in total private sector lending. *Public Bank*: = 1 if bank is publicly owned. *Share of Business Lending*: weight of lending to firms in total lending. *Leverage*: assets over equity. *Liquidity*: ratio of liquid assets over total assets. *ROE*: return on equity. *Average Loan Rate*: average interest rate on local currency loans. *Average Deposit Rate*: average interest rate on local currency deposits. *ROA*: return on assets. *Assets*: total assets. *Share of Aggregate Deposits*: participation of each bank in aggregate deposits.

More importantly, Figure 2.3 illustrates the relationship between bank exposure and credit supply using raw data. Panel (a) aggregates firm-level debt within each exposure quartile, while panel (b) restricts the sample to firms borrowing from both low- and high-exposure banks, allowing for within-firm comparisons across banks. In both panels, lending to firms follows parallel trends across exposure groups in the two years prior to the policy, indicating the absence of pretrends or anticipation effects. Following the policy, even without controlling for fixed effects, banks with higher exposure exhibit a clear increase in total lending.

Intuitively, the policy was introduced as a one-off fiscal measure in a context unrelated to banking sector distress, which reduces the likelihood of observing pretrends or anticipation effects.



(A) Across Banks.



(B) Within Firms & Across Banks.

Figure 2.3. Above *vs* Below Median Bank Exposure - Total Credit.

Note: In Panel (a) firm level debt is added up for each exposure quantile. Panel (b) restricts the sample to firms that had debt with banks both above and below median exposure in $2016Q2$. We construct the evolution of their real debt from each of these group of banks. In other words, firms appear in both the red and blue lines. In both cases the base period is $2016q2$ and total credit is expressed in current US dollars.

Firm-level Exposure

We measure firm-level exposure as:

$$S_i = \sum_b \omega^0_{ib} S_b, \tag{2.2}$$

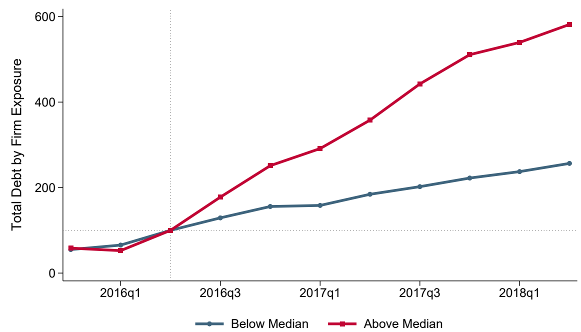
where S_i is firm's i exposure to the tax amnesty shock and ω^0_{ib} is the weight of bank b on firm's i total borrowing in the quarter preceding the tax amnesty. The exposure of firm i is a weighted average of the exposure of the banks from which firm i borrowed at the start of the tax amnesty. We argue that the fact that different firms are connected to banks with different degrees of exposure provides a source of plausibly exogenous variation in the credit supply faced by individual firms.

As an initial overview of the results—which we will explore in more detail in Section 2.4.3 —Figure 2.4 provides information on the debt dynamics of firms with different levels of exposure over time, as observed in the raw data. The left panel illustrates the evolution of total credit for firms with below- and above-median exposure, while the right panel depicts the evolution of dollar-denominated debt. Both measures are expressed in current US dollars and relative to $2019q2$ to facilitate comparison.

First, the raw data do not show evidence of differential pretrends across firms with different exposure to the shock. Second, after tax amnesty, firms with higher exposure experienced stronger total debt growth, particularly in dollar-denominated debt, relative to their lower exposure counterparts.

We will revisit these results in Section 2.4.3 when we formally analyze the impact of the tax amnesty on firm outcomes.

(A) Total Debt (in US Dollars).



(B) Dollar Debt (in US Dollars).

Figure 2.4. Below *vs.* Above Median Firm Exposure.

Note: Panel A shows the evolution of total debt, expressed in current US dollars, for firms below and above median exposure. The period covered is $2014q1-2018q2$. Panel B plots the evolution of dollar debt, expressed in current US dollars, for each group of firms. Panel B covers the period $2015q4-2018q2$ for which dis-aggregated data by firm-currency is available.

2.4.2 Bank Lending Channel

We begin by testing the presence of a bank lending channel in response to the funding shock generated by the tax amnesty. Specifically, we examine whether banks that were more exposed to the inflow of tax amnesty funds extended more credit than less-exposed banks. To isolate the effects of credit supply from potential demand-side confounders, we follow the within-firm estimator proposed by Khwaja and Mian (2008). This estimator exploits the fact that many firms borrow simultaneously from multiple banks and compares lending changes by differently exposed banks to the same firm, effectively controlling for any time-varying firm-specific shocks to credit demand.

Assume that there are only two time periods $t = 0, 1$, before and after the tax amnesty shock. Then, we can express the level of debt of the firm i with bank b in period t as:

$$Y_{ib0} = \alpha_{i0} + \nu_{ib0} + \Gamma \mathbb{X}_{b0} + \varepsilon_{ib0} \tag{2.3}$$

$$Y_{ib1} = \alpha_{i1} + \nu_{ib1} + \Gamma \mathbb{X}_{b1} + \beta S_b + \varepsilon_{ib1}. \tag{2.4}$$

The realization of Y_{ibt} depends on a number of objects. First, α_{it} captures firm-time credit demand shocks that a firm spreads homogeneously between all its lenders. Second, ν_{ibt} are credit demand shocks specific to a certain bank-firm relationship. For example, a firm needs a type of credit line that is offered only by a specific bank. $\mathbb{X}_{bt}$ captures all other bank-specific liquidity shocks except for the tax amnesty shock. Finally, S_b is our measure of bank-level exposure to the tax amnesty shock and ε_{ibt} is an idiosyncratic shock. Taking first differences yields:

$$\Delta Y_{ib} = \Delta\alpha_i + \Delta\nu_{ib} + \Gamma \Delta\mathbb{X}_b + \beta S_b + \tilde{e}_{ib}. \tag{2.5}$$

We are interested in estimating β, which captures the marginal effect of an increase in the exposure of bank b on its lending to firm i. β has a causal interpretation under the condition that bank exposure is not correlated with other factors that affect changes in credit demand and bank liquidity. We take equation 2.5 to the data as follows:

$$\log L_{ibt} = \alpha_{it} + \nu_{ib} + \beta \; S_b \times Post + \Gamma \mathbb{X}_{b,pre} \times Post + e_{ibt}. \tag{2.6}$$

In order to make *within-firm* comparisons we restrict the analysis to multi-bank firms. This allows us to incorporate firm-time fixed effects, α_{it}, which absorb time-varying firm-specific credit demand shocks that firms spread equally across all their lenders. We also include *firm-bank* fixed effects to control for aspects specific to each firm-bank match. $Post$ is a dummy equal to one after the start of the tax amnesty. Lastly, we incorporate as bank-level controls a set of bank characteristics, measured at the quarter prior to the shock, and we interact them with our $Post$ dummy. Based on the balance tests from Section 2.4.1, we include bank ownership status (private vs. public) and bank size as controls in our preferred specification.

Results

Table 2.1 presents our results. Point estimates measure the percent increase in lending to firm i by bank b for every percentage point (p.p.) increase in bank exposure. The results in column (1) include only firm-time fixed effects. Column (2) adds bank-firm fixed effects and bank ownership controls. Our preferred specification, in column (3), also adds bank size controls. Comparing between banks in the 75^{th} *versus* 25^{th} percentile of exposure[9], we find an increase in lending of 10.7% .

In column (4) we allow for the effect of exposure to differ across exporters and non-exporters, where the exporting status of a firm is defined at baseline. Exporters are the main recipients of dollar bank credit [10] thus we would expect a higher impact on credit supplied to them. The coefficient on the interaction term is positive and statistically significant. If we compare between banks in the 75^{th} *versus* 25^{th} percentile of exposure, we find an increase in lending to the *same* exporting firm of 18.4%. The same comparison for non-exporting firms yields a 7% increase in lending. In summary, we find evidence of a bank lending channel in response to the tax amnesty shock. Banks that were more exposed to the tax amnesty granted more credit relative to their less exposed counterparts. The effect was heterogeneous between lending to different types of firms. The impact on exporters was greater and more than double that of non-exporters.

[9] The difference in exposure between the 75th and 25th percentiles of exposure is 10 percentage points.

[10] In Appendix B.2 we provide details on the distribution of dollar-denominated bank credit to firms at the start of the tax amnesty.

Table 2.1. Bank Lending Channel.

	(1)	(2)	(3)	(4)
Bank Exposure × Post	0.21**	1.17***	1.07***	0.71***
	(0.097)	(0.125)	(0.130)	(0.141)
... × Exporter				1.13***
				(0.395)
N	1655983	1655981	1655981	1655981
Firm-Time FE	Yes	Yes	Yes	Yes
Bank-Firm FE	No	Yes	Yes	Yes
Public Bank-Time FE	No	Yes	Yes	Yes
Bank Size - Time FE	No	No	Yes	Yes
R-squared	0.61	0.87	0.87	0.87

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the bank-firm level. Point estimates measure the percent increase in L_{ibt} for every $1 p.p.$ increase in bank exposure, S_b. $Post$ is equal to one between $(2016Q3, 2018Q2)$. The sample consists of all firms that have debt with two or more banks. We winsorize the top and bottom 1% of observations to limit the influence of outliers. Bank size is measured as the participation of each bank in total deposits at baseline and public bank is an indicator for a bank being publicly owned.

Identification Assumptions

Our empirical strategy for the bank lending channel exploits variation in the intensity of the liquidity shock across banks lending to the same firm, following Khwaja and Mian (2008).

The key identification assumption is that, conditional on our set of fixed effects and controls, banks that received more amnesty funds were not differentially exposed to other unobserved shocks that affected either (i) their lending capacity (ii) or the credit demand they faced. Under this assumption, differences in post-amnesty lending behavior across banks capture a causal supply-side response to the inflow of funds.

To structure our identification discussion, we distinguish two main potential threats, one arising from unobserved bank-level supply shocks and the other from firm-level credit demand shocks.

On the bank supply side, a potential threat to identification would arise if unobserved funding shocks or shifts in banks' lending capacities were correlated with the inflow of amnesty deposits. We mitigate this concern in two ways. First, we show that banks were balanced in terms of observable pre-amnesty characteristics across levels of amnesty exposure, suggesting that differential deposit inflows were not systematically related to pre-existing differences in bank health or lending capacity. Second, we show that our results remain unchanged when adding controls for observable bank characteristics—such as size and ownership status—interacted with post-amnesty time indicators, suggesting that observable proxies for bank-level funding conditions do not account for the estimated effects.

On the credit demand side, our identification assumption further requires that changes in credit demand faced by banks are not systematically correlated with their exposure to the shock. To mitigate concerns about firm-level demand shocks, we exploit within-firm variation, following the methodology developed by Khwaja and Mian (2008). Specifically, by restricting our analysis to firms borrowing simultaneously from multiple banks with varying levels of exposure, we effectively absorb any firm-level shocks affecting their overall demand for credit. Panel (b) of Figure 2.3 illustrates this clearly, showing the differential credit growth for highly exposed versus less exposed banks using only firms that borrowed from both groups prior to the amnesty. Because each firm contributes to both high and low exposure bank groups, any observed differences post-amnesty cannot be attributed to aggregate firm-level demand shocks.

A remaining threat to our identification arises if there are bank-firm specific credit demand shocks, that is, shocks affecting certain bank-firm pairs disproportionately due to factors such as bank specialization or differential exposure to other simultaneous economic shocks. A particularly relevant potential shock relates to foreign trade financing. The contemporaneous trade liberalization that occurred in Argentina between 2016 and 2017 could have led international trade-oriented firms (exporters or importers) to redirect their credit demand specifically towards banks specialized in trade finance. If these banks were systematically more or less exposed to the amnesty-induced inflow, our estimates would be biased.

In Appendix B.4, we show results for a number of alternative specifications. First, we present results controlling for different measures of bank specialization to foreign trade financing and show that our findings remain largely unchanged. Second, we show that our results remain unchanged when adding, one at a time, additional bank-level controls to our preferred specification. Each of these robustness exercises is designed to address the identification threats identified above.

2.4.3 Firm-level Outcomes

In the previous section, we established that firms borrowed relatively more from banks that were more exposed to the tax amnesty shock. But did this relative shift towards more exposed banks result in more total borrowing by firms? Or did firms mostly reallocate their debt portfolio from less exposed towards more exposed banks? This section tackles this question, which is in turn key to understanding the potential effects of the tax amnesty shock on firms' real performance.

We begin by describing our estimation strategy. For the remainder of this section, we will be interested in measuring the effect that firm-level exposure, S_i, had on a series of firm-level outcomes Y_{it} - including total credit, access to dollar credit markets, trade performance and employment growth. For simplicity, suppose that there were only two periods $t = 0, 1$ that correspond to before and after the tax amnesty, respectively. We can think of outcome Y_{it} as follows:

$$Y_{i0} = \alpha_i + \eta_{i0} + \Gamma \mathbb{X}_{b(i)0} + \varepsilon_{i0}, \tag{2.7}$$

$$Y_{i1} = \alpha_i + \eta_{i1} + \Gamma \mathbb{X}_{b(i)1} + \beta_2 S_i + \varepsilon_{i1}. \tag{2.8}$$

The realization of outcome Y_i in period $t = 0$ is the result of (i) a time-invariant firm shifter, α_i; (ii) a time-varying firm-level shifter, η_{it}; (iii) a credit supply shifter, $\mathbb{X}_{b(i)0}$, coming from the banks from which firm i sources its debt (other than the tax amnesty); plus, (iv) an idiosyncratic component ε_{i0}. In period $t = 1$, we add the credit supply shock resulting from the tax amnesty, S_i. Taking first differences yields:

$$\Delta Y_i = \Delta \eta_i + \beta_2 S_i + \Gamma \Delta \mathbb{X}_{b(i)} + \tilde{e}. \tag{2.9}$$

Our main object of interest is β_2. It captures the causal effect of firm exposure on outcome Y_i. To identify β_2, two conditions must be satisfied. First, there should be no systematic correlation between changes in firm-level shifters and firm exposure to the shock. Formally, we need $E[\Delta\eta_i S_i] = 0$. For example, if Y_{it} is firm-level credit, then credit demand shocks hitting specific firms over time should not be related to their exposure to the shock. The second condition speaks to the correlation between firm exposure and, what can be thought of as, all other liquidity shocks hitting the banks from which firm i borrows. Formally, we need $E[\mathbb{X}_{b(i),t} S_i] = 0$. This would be challenged if more exposed banks were also more likely to be hit by other funding shocks than less exposed banks over time.

We take equation 2.9 to the data using the following difference-in-difference specification:

$$\log Y_{it} = \beta_2 \, S_i \times Post + \Gamma \mathbb{X}_{b(i)} \times Post + \gamma_i + \gamma_{jt} + e_t, \tag{2.10}$$

where the dependent variable is the logarithm of the outcome of interest (unless otherwise noted) in period t, $Post$ is a dummy equal to 1 after the start of the tax amnesty, $\mathbb{X}_{b(i)}$ is a vector of firm-level weighted average bank controls, γ_i are firm fixed effects and γ_{jt} capture industry-time fixed effects. Firm-level bank controls are computed as weighted averages of baseline bank characteristics, where the weight is the share of each bank b in firms i debt at baseline. We study the time period between 2014-2018. The frequency of analysis is either quarterly or annual, depending on the outcome variable.

We address the concern that firm exposure could potentially be correlated with time-changing shifters at the industry level, such as credit lines for specific industries, by adding time-industry fixed effects. These take care of any trends in firm-level shifters that are shared within narrowly defined industries (6-digit level). We take a conservative stance and include the same two bank controls as in our *within-firm* regressions and interact them with the $Post$ dummy. These bank controls are ownership status (privately or publicly owned bank) and bank size, which we measure as a bank's participation in total deposits.

Lastly, we estimate the event-study analogue of Equation 2.10 to check for pretrends and to study the dynamic impact of the shock. Our baseline dynamic difference-in-difference specification is:

$$\log Y_{it} = \sum_{s\in(\pm h)} \left(\beta_s S_i + \Gamma_s \mathbb{X}_{b(i)}\right) \times \mathbb{1}(t=s) + \gamma_i + \gamma_{jt} + e_{it}, \tag{2.11}$$

where $\mathbb{1}(t=s)$ are quarterly or yearly time dummies and h is the horizon of analysis (e.g., eight quarters before and after the shock). The coefficient β_s measures the s-period cumulative effect of firm exposure on outcome Y. The absence of pre-trends requires $\hat{\beta}_s = 0 \;\forall\; s \leq 0$.

Identification Assumptions

Our identification assumption is that, conditional on our set of fixed effects and controls, firms more exposed to the tax amnesty shock, through their pre-existing borrowing relationships with more affected banks, were not differentially affected by unobserved shocks to our outcomes of interest. We rely on the standard parallel trends assumption: that in the absence of the amnesty, outcomes for more and less exposed firms would have followed similar trends within the same industry and foreign trade participation group.

The inclusion of industry- foreign trade participation-time fixed effects is central to this strategy, as it absorbs any time-varying shocks common to firms operating in the same product markets and sharing the same exposure to the contemporaneous trade liberalization. Reassuringly, none of the outcome variables display significant pre-trends, and our quarterly firm-level loan data allow us to verify that the amnesty began affecting firms precisely in mid-2016, mitigating concerns that differential exposure reflects other contemporaneous shocks rather than the amnesty itself.

Firm Borrowing

In this section we explore the impact of the tax amnesty shock on credit growth and credit access through three different avenues. First, we study whether the positive effects identified in the within-firm regression extend to the overall borrowing behavior of firms. This aims to determine if firms that were more exposed to the shock experienced higher credit growth - a mechanism often referred to as the *firm borrowing channel*. Second, we evaluate the performance of the dollar credit market, as the tax amnesty shock significantly enhanced banks' capacity to issue dollar-denominated credit. The well-functioning of foreign-currency, in particular dollars, credit markets has been pointed out as an important driver of firm performance, specially in emerging economies. Thus, we evaluate the tax amnesty's effectiveness in broadening and deepening the domestic market for dollar credit.

Finally, we examine how the tax amnesty influenced the currency composition of firms' borrowing, shifting our focus from aggregate credit growth to the currency choices made by borrowing firms. By leveraging the currency-specific nature of the shock, we estimate the reduced-form effect of dollar liquidity expansions on the share of dollar-denominated debt in firms' credit portfolios. As we show in the model in Section 2.5, this reduced-form elasticity is informative about the structural elasticity of substitution between peso- and dollar-denominated debt, providing a moment that helps identify this parameter.

Total Credit

We begin by analyzing the effect on firms' total borrowing. Note that this analysis allows to capture changes in total credit at the firm level and therefore directly speaks to credit expansion, rather than substitution across banks. As such, it complements the bank-lending regressions with firm-year fixed effects, which may be reflecting a reallocation of credit across banks with different exposure. By focusing on total firm borrowing, this approach allows us to assess whether firms actually increase overall credit, rather than simply experiencing a reshuffling of credit across lenders.

We employ quarterly data for single and multi-bank firms[11], winsorize the dependent

[11] The *bank-lending* regressions restrict the sample to multi-bank firms and pre-existing bank relations. Firm-level regressions consider both single and multibank firms, plus all credit available to the firm at each point of time. Thus, total credit incorporates both preexisting bank relations and newly formed ones.

variable at the 1% level and cluster standard errors at the firm level. Reported coefficients measure the percentage increase in total credit for every additional percentage point in firm exposure.

Table 2.2 shows the results of running specification (2.10) with total firm credit as dependent variable. Column (1) only includes firm fixed effects and column (2) adds (6-digit) industry-time fixed effects. Both specifications show positive and statistically significant results. Column (3) shows the results from our preferred specification which includes firm fixed effects, (6-digit) industry-time fixed effects and firm-level bank controls. For the average firm, the effect of an additional percentage point in exposure is .97%. When comparing between firms in the 75^{th} *versus* 25^{th} percentiles of exposure[12], our estimates imply an increase of 9.7% in total credit. The magnitude of this effect is almost the same as that found for the *within-firm* exercise. We interpret this finding as evidence that exposure to the tax amnesty shock primarily influenced total firm credit, rather than firms reallocating existing borrowing across differentially exposed banks. In column (4) we add an interaction term with a firm's exporting status. We find that the effect of exposure was stronger among exporting firms, which is consistent with our *within-firm* findings.

We complement the above results by estimating the dynamic difference-in-difference specification outlined in equation (2.11). We set the base period to 2016q2 which corresponds to the quarter preceding the launch of the tax amnesty. The event study plot in Figure 2.5 shows no significant evidence of pretrends for at least a year and a half prior to the shock. The plot shows that the positive effects of firm exposure gradually build up over time with a peak effect of around 2% at the 2-year horizon.

The above results establish that firms that were more exposed to the tax amnesty shock, due to their pre-existing bank relations, experienced higher credit growth.

[12] The difference in exposure between a firm in the 75h percentile of exposure and one in the 25th percentile is 10 percentage points.

Table 2.2. Firm Borrowing - Total Credit.

	(1)	(2)	(3)	(4)
Firm Exposure × Post	1.39***	1.21***	0.97***	0.86***
	(0.093)	(0.105)	(0.141)	(0.144)
... × Exporter				2.17***
				(0.613)
N	1362654	1355733	1355733	1358139
Firm FE	Yes	Yes	Yes	Yes
Industry-Time FE	No	Yes	Yes	Yes
Public Bank-Time FE	No	No	Yes	Yes
Bank Size - Time FE	No	No	Yes	Yes
Number of Firms	75879	75494	75494	75628
R-squared	0.85	0.85	0.85	0.85

Note: ** $p < 0.05$, *** $p < 0.01$. (6-digit) Industry-time fixed effects are allowed to differ across exporters and non-exporters. Bank size and bank ownership controls are firm-level averages. Nominal variables are winsorized at the top and bottom 1% to limit the influence of outliers. The time period spans between $2014q2 - 2018q2$ and the frequency of the data is quarterly. *Post* indicates the period $(2016q3, 2018q2)$. Point estimates measure the percent increase in total credit for every percentage point of firm exposure. The standard deviation of firm exposure is 10 percentage points. In column 4, the exporting status of a firm is defined at baseline.

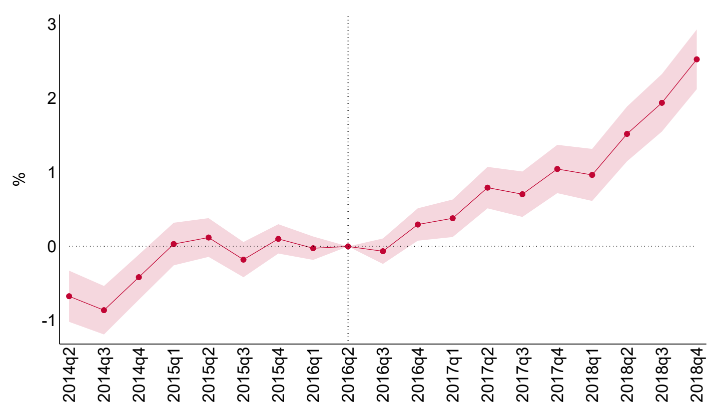


Figure 2.5. Total Firm Borrowing.

Note: We report 95% confidence intervals. Standard errors are clustered at the firm level. We include firm fixed effects, (6-digit) industry-time fixed effects and firm level averages of bank size and bank ownership interacted with time dummies as additional controls. The base period corresponds to 2016q2. Nominal variables are winsorized at the top and bottom 1% to limit the influence of outliers. The plot shows the estimated coefficients on firm exposure for each quarter. Point estimates measure the percentage increase in total debt for every 1 percentage point increase in firm exposure relative to baseline. The standard deviation of firm exposure is 10 percentage points.

Access to Dollar Credit Markets

The tax amnesty led to an increase in the foreign currency liquidity of banks, thus potentially allowing them to grant more dollar credit to domestic borrowers. The role of access to foreign currency credit markets for firm performance has received significant attention in corporate finance and international economics literature, particularly following the freezing of international markets during the Great Recession. Fluctuations in dollar credit availability are linked to reduced firm investment (e.g. Kalemli-Ozcan et al. (2016),Elias (2021)), diminished export dynamism (e.g. Paravisini et al. (2015c)), and weaker employment growth performance (e.g. Hardy (2023)). The tax amnesty shock resembles the opposite of a reversal in international capital flows, providing a unique opportunity to study how changes in foreign currency funding affect firms' participation in foreign-currency credit markets.

To examine the impact of the tax amnesty on firms' access to domestic dollar-denominated credit, we leverage that our dataset includes the loan's currency denomination from 2015q4 onward. The currency denomination of a bank liquidity shock theoretically increases loanable funds in both foreign and domestic currencies. However, due to currency mismatch regulations, domestic banks are restricted in how they can use dollar-denominated liabilities. Specifically, they must match dollar liabilities with dollar assets. This requirement creates a direct link between the currency denomination of the shock and the supply of dollar credit.

In addition, access to domestic dollar credit is restricted to firms whose revenues are, at least partially, linked to the evolution of the dollar exchange rate. These firms are primarily exporters, suppliers of exporters, or producers of highly tradable goods whose prices are set internationally. At the beginning of 2016, 27% of the firms in our sample had some amount of dollar-denominated bank credit. This percentage drops to 13% when considering only firms with an outstanding balance of at least 500[13] and increases to 50% when conditioned on the firm being an exporter and/or importer. Regarding the distribution of dollar credit, exporting firms held around 80% of total dollar credit at the start of the sample, while exporters and importers

[13]This restriction excludes firms that use, for example, credit cards abroad, etc.

combined accounted for almost 90%. Appendix B.2 provides further details on the distribution of domestic dollar credit before and after the tax amnesty.

As pointed out above, a small number of firms had dollar debt before the tax amnesty which restricts the power of intensive margin analysis. For this reason, we focus our analysis on the extensive margin effects of the tax amnesty. This allows us to study changes in access to dollar credit for a larger sample of firms. We define two outcomes related to the ability of firms to borrow in dollars. First, we construct an indicator variable that is equal to one if the firm i has positive dollar credit in period t, and zero otherwise. This variable allows us to estimate the effect of the tax amnesty on the probability of having access to dollar credit for the average firm. Second, we construct an indicator equal to one in every quarter in which the firm i experienced an increase in dollar credit, and zero otherwise. This outcome variable allows us to study a combination of intensive and extensive margin results without dropping firms with zero outstanding balances, giving a more general picture of the response of dollar credit growth for the average firm.

Table 2.3 summarizes our results. Columns (1) and (2) use the probability of having dollar credit as the dependent variable. Our results indicate that firm exposure had a positive and significant effect on this probability. Our preferred specification in column (2) finds that an additional percentage point of exposure increased it by .3 percentage points. Comparing between firms in the 75^{th} *versus* 25^{th} percentiles of exposure, we find an increase of 3 percentage points in the probability of borrowing in dollars. This represents a 10% increase relative to the unconditional probability of holding dollar debt at the start of the sample. In columns (3) and (4) we show results for our dollar credit growth indicator. We estimate that an additional percentage point of exposure increased the probability of dollar credit growth by .22 percentage points. When comparing between firms in the 75^{th} *versus* 25^{th} percentiles of exposure, this is an increase of 2.2 percentage points in our outcome variable.

Table 2.3. Access to Dollar Credit Market.

	(1) Dollar Debt	(2) Dollar Debt	(3) Dollar Debt Growth	(4) Dollar Debt Growth
Firm Exposure × Post	0.13***	0.31***	0.09***	0.22***
	(0.021)	(0.028)	(0.017)	(0.023)
N	832009	832009	832009	832009
Firm FE	Yes	Yes	Yes	Yes
Industry-Time FE	Yes	Yes	Yes	Yes
Public Bank-Time FE	No	Yes	No	Yes
Bank Size - Time FE	No	Yes	No	Yes
Number of Firms	75649	75649	75649	75649
R-squared	0.65	0.65	0.34	0.34

Note: Standard errors in parentheses and clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (6-digit) Industry-time fixed effects are allowed to differ across exporters and non-exporters. Bank size and bank ownership controls are firm-level averages. The time period spans between $2015Q4 - 2018Q2$. $Post$ indicates the period $(2016q3, 2018q2)$. Point estimates measure the $p.p.$ increase in the dependent variable for every 1 $p.p.$ of firm exposure.

Firm-level Currency Choices

Next, we study how the tax amnesty shock affected the currency choices made by firms. Specifically, we exploit the currency-specific nature of the shock and institutional features of Argentina's financial system to estimate the reduced-form effect of a dollar liquidity expansion on the share of dollar-denominated debt in firms' credit portfolios. As we show in the model, this reduced-form elasticity is informative about the structural elasticity of substitution between peso and dollar loans, which we denote κ:

$$\kappa = \frac{\partial \ln L^{Pesos}}{\partial \ln L^{*}}, \tag{2.12}$$

where the superscript * denotes a dollar-denominated quantity.

Emerging markets are recurrently subject to fluctuations in the availability and access to foreign-currency credit markets. For example, the global financial cycle is well documented in papers such as Obstfeld and Zhou (2022), Bruno and Shin (2023), Rey (2014), Miranda-Agrippino and Rey (2022), Boyarchenko and Elias (2024), among others. The real and aggregate effects of relative fluctuations in domestic versus foreign currency funding will partly depend on firms' ability to substitute one source of funds for the other. If both types of financing serve the same purposes or can be used to finance the same types of activities, then we would expect peso and dollar credit to act as substitutes for each other. However, different forms of financing may be better suited for different productive tasks based, for example, on the maturity of the project or the source of revenues. This can lead to complementarities between debt in different currencies, thereby limiting the ability of firms to hedge against fluctuations in dollar credit by turning to peso markets. Although there is little evidence on how these dynamic portfolio choices take place, knowledge of the substitution parameter is crucial for understanding the aggregate effects of fluctuations in the availability of foreign currency debt.

To estimate the reduced-form effect of dollar credit on currency composition, our cross-sectional design requires a shock that shifts the supply of dollar credit heterogeneously across firms, while any induced change in peso credit supply is orthogonal to firm-level exposure S_i. This is weaker than requiring peso supply to be unaffected by the amnesty altogether: it permits a uniform expansion in peso lending capacity, for instance through system-wide portfolio rebalancing or freed up lending limits, as long as this expansion does not favor firms with higher dollar exposure. We argue that the tax amnesty shock satisfies this condition. The regulatory requirements that were in place at the time tightly limited domestic banks' on- and off-balance sheet currency exposure, ruling out a direct channel through which dollar inflows could be converted into additional peso lending. Any indirect effect on peso supply, such as firms freeing up peso borrowing capacity as they shift toward dollars, would have to operate through bank-firm specific channels correlated with S_i to threaten identification. We show in Appendix B.5 that bank-level exposure has no detectable effect on the evolution of average peso lending rates, suggesting that any such indirect channel is not quantitatively important.

We estimate specification (2.10) with the share of dollar debt as the dependent variable. Our preferred specification includes the same set of fixed effects and firm-level bank controls as for the firm borrowing channel. We also include an interaction between the baseline share of dollar debt and the *Post* treatment dummy to restrict our comparisons to firms with similar currency composition at baseline but heterogeneous in their exposure to the tax amnesty shock.

Table 2.4. Share of Dollar Debt.

Y_{ft}: Share of dollar credit	(1)	(2)	(3)
Firm Exposure × Post	.12***	0.08***	
	(0.007)	(0.01)	
... × Non-Exporter			0.08***
			(0.01)
... × Exporter			.14**
			(0.07)
α_f, α_{jzt}	Yes	Yes	Yes
X_f controls	No	Yes	Yes
Number of Firms	73795	73795	73795
Baseline dollar share - All firms	2%		
Baseline dollar share - Exporters	7%		

Note: Standard errors in parentheses and clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (6-digit) Industry-time fixed effects are allowed to differ across exporters and non-exporters. Bank size and bank ownership controls are firm-level averages. The time period spans between $2015Q4 - 2018Q2$ and the frequency of the data is quarterly. *Post* indicates the period $(2016q3, 2018q2)$. Point estimates measure the percentage point increase in the share of dollar credit for every percentage point of firm exposure. The standard deviation of firm exposure is 10 percentage points. In column 3, the exporting status of a firm is defined at baseline.

Table 2.4 presents our results. We estimate that the average firm experiences an increase in the participation of dollar in total borrowing of .12 percentage points for every percentage point of exposure. Moving from the 25th to the 75th percentile of exposure implies an increase in the dollar share of 1.2 percentage points. This represents a 50% increase relative to the baseline share of dollar debt of the average firm. Column 3 shows that the point estimate for exporters is almost double that for non-exporters, however, the effects are less precisely estimated. In Section 2.5 we show how this reduced-form elasticity maps to the structural elasticity of substitution between peso and dollar debt: under the identifying assumptions of our cross-sectional design, a larger currency-composition response to the dollar shock is informative of higher substitutability between the two forms of credit.

Imports

We study how the tax amnesty shock affected the importing behavior of firms. Our hypothesis is that improved access to credit, particularly dollar-denominated bank loans, helps credit-constrained firms overcome the sunk costs of discovering foreign suppliers, as well as finance international purchases.

We restrict the analysis to firms that were already importing before the amnesty and focus on annual imports. We then examine two margins of adjustment: (i) the intensive margin, which focuses on import volumes, and (ii) the extensive margin, focusing on the number of products and the number of origins.

The results of our static difference-in-difference are presented in Table 2.5. Just as in the analysis of firm credit outcomes, we include firm fixed effects, industry-time fixed effects and the same set of firm-level weighted average bank controls. We measure the number of imported products as the number of distinct products imported by a firm during each year. The number of origins is computed analogously.

Table 2.5. Effect on Imports.

	(1) Total Imports	(2) Intermediates	(3) Capital Goods	(4) Origins	(5) Products
Firm Exposure × Post	1.27***	1.23***	0.44*	0.10	0.11
	(0.315)	(0.357)	(0.248)	(0.088)	(0.142)
N	57079	48875	33023	57079	57079
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
Public Bank-Year FE	Yes	Yes	Yes	Yes	Yes
Bank Size-Year FE	Yes	Yes	Yes	Yes	Yes
Number of Firms	13489	11850	8660	13489	13489
R-squared	0.81	0.84	0.73	0.88	0.85

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the firm level. Point estimates in columns $1-3$ measure the percent increase in the outcome of interest for every 1 $p.p.$ increase in firm exposure. Point estimates in columns $4-5$ are expressed in $p.p.$. The sample consists of all firms that were importing at the start of the tax amnesty. Nominal variables are winsorized the top and bottom 1% to limit the influence of outliers.

We find that higher firm exposure resulted in an increase in their total imports relative to less exposed firms. If we compare between firms in the 75^{th} *versus* 25^{th} percentile of exposure, we find an increase of 12.7% in total imports. We then look at different types of imported products and find evidence of an increase in both intermediate and capital goods. This is suggestive that firms were using the newly funds to invest. Lastly, we find no evidence that greater exposure to the tax amnesty affected the extensive margin, either through the introduction of new products or sourcing from new origins.

Exports

Next, we examine how increased access to credit affects export performance. Since exporting is a finance-intensive activity, looser dollar-credit conditions can influence exporting by easing borrowing constraints. An increase in exports can also be indicative of a firm increasing its productivity. Table 2.6 reports the main results. We restrict the sample to firms that were exporting before the tax amnesty. As in previous exercises, we include firm fixed effects, year fixed effects, and the same set of firm-level weighted average bank controls. We replace industry-time fixed effects with main exported product (6-digit HS) by time fixed effects to better account for time-varying demand shocks specific to the good being sold. Columns (1)–(3) sequentially add fixed effects and controls.

Throughout the specifications, we observe a positive effect on exports. Our preferred specification in column (3) shows a positive and significant effect of exposure on total exports. A comparison between firms at the 75th and 25th percentiles of exposure implies a 13.7 percent increase in exports.

We then study whether the effect of firm exposure was homogeneous between exporters of different sizes. If the increase in credit supply operates through a relaxation in firm's borrowing constraints, then we should see a greater impact for firms closer to this constraint or with less access to credit markets. In general, those are small- and medium-sized firms. To evaluate whether this is the case, we estimate the impact of firm exposure for each quartile of the exporting

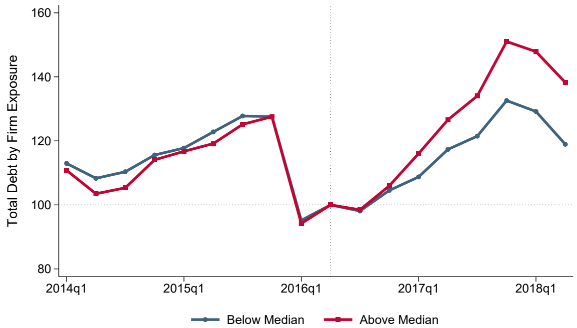
Table 2.6. Effect on Exports.

	(1)	(2)	(3)	(4)
Firm Exposure × Post	0.99**	1.10**	1.37**	
	(0.41)	(0.43)	(0.58)	
... × Exports Quantile = 1				4.04***
				(1.19)
... × Exports Quantile = 2				1.02
				(0.95)
... × Exports Quantile = 3				1.02
				(0.92)
... × Exports Quantile = 4				-0.09
				(0.87)
N	19075	19075	19075	19075
Firm FE	Yes	Yes	Yes	Yes
Product×Year FE	Yes	No	No	No
Product×Year×Destination FE	No	Yes	Yes	Yes
Bank Controls	No	No	Yes	Yes
Number of Firms	4,925	4,925	4,925	4,925
R-squared	0.92	0.92	0.93	0.93

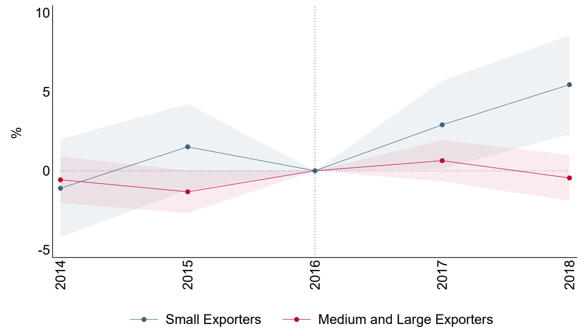
Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the firm level. Point estimates measure the percent increase in export values for every 1 p.p. increase in firm exposure. We restrict the sample to all firms that were exporting at the start of the tax amnesty. We winsorize the top and bottom 1% of observations to limit the influence of outliers.

distribution. We report our findings in column (4). In fact, we find strong evidence that the increase in exports was driven mainly by smaller exporters (in terms of their exported values). If we compare between a highly exposed exporter and a low exposed exporter, both in the lower quartile of the export distribution, we find an increase of 40% in total exports. The estimated coefficients are positive and smaller for the two middle quartiles but statistically insignificant. Finally, the impact on the top quartile of exporters is null. These results are consistent with the fact that smaller firms tend to be more financially constrained. In addition,

Our dynamic difference and difference plots in Figure 2.6 complement these findings.



(A) Full Sample.



(B) Heterogeneity by Exporter Size.

Figure 2.6. Exports.

Note: We report 95% confidence intervals. Standard errors are clustered at the firm level. The plot shows the estimated coefficients on firm exposure for each year. Point estimates measure the percentage increase in export values for every 1 p.p. increase in firm exposure. We restrict the sample to all firms that were exporting at the start of the tax amnesty and winsorize the top and bottom 1% of observations to limit the influence of outliers.

Labor Outcomes

Tax amnesties and inflows of capital have been studied for their potential effects on government revenue and capital flows, but a lesser-known potential outcome is the indirect benefits they may offer to employers and workers of firms connected with banks exposed to these capitals. By overcoming liquidity constraints, firms can increase investment and productivity, which can lead to an increase in the number of employees or the wages paid. In other words, the benefits of liquidity shocks stemming from tax amnesties can trickle down to workers. Although this effect is often overlooked, we shed light on this potential outcome and provide an analysis of the impact of tax amnesties on labor outcomes.

Table 2.7. Employment and Wages.

	(1) Employment	(2) Employment	(3) Employment	(4) Wages	(5) Wages	(6) Wages
Firm Exposure × Post	0.20***	0.17***	0.17***	0.01	0.02	0.02
	(0.03)	(0.04)	(0.05)	(0.02)	(0.03)	(0.03)
... × Post			0.08			0.06
			(0.15)			(0.09)
N	1258952	1258952	1258952	1258260	1258260	1258260
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Time-Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Public Bank-Time FE	No	Yes	Yes	No	Yes	Yes
Bank Size - Time FE	No	Yes	Yes	No	Yes	Yes
Number of Firms	69828	69828	69828	69812	69812	69812
R-squared	0.96	0.96	0.96	0.85	0.85	0.85

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Standard errors in parentheses and clustered at the firm level. Point estimates measure the percent increase in the outcome of interest for every 1 p.p. increase in firm exposure. Employment refers to the total number of formal employees in each quarter t. Wages corresponds to the average nominal wage paid by each firm in quarter t. Bank controls include bank size and bank ownership status interacted with the $Post$ indicator as in previous regressions.

Table 2.7 reports the results for our difference-in-difference specification. The first two columns show results on employment without and with bank controls, respectively. The third column incorporates an interaction term for the export status. Columns (4)-(5) show results for firm-level average wages without and with bank controls, respectively, and column (6) adds an interaction term for export status. We find positive and significant effects of firm exposure on employment. The export status term is also positive, but statistically insignificant. Our findings imply that an average firm in the 75^{th} percentile of exposure experienced an average increase of 2.1% in employment relative to the average firm in the 25^{th} percentile, during the aftermath of the tax amnesty. In contrast, the impact of firm exposure to the tax amnesty on average wages is not statistically significant. In a context of increasing employment, any increase or decrease in the wages of workers who remain with the firm could have been offset by changes in the composition of the workforce.

Lastly, Figure 2.7 shows the event study for firms' employment. Reassuringly, we observe no pretrends for these firms in the years before the event and a substantial and persistent increase in employment after the tax amnesty.

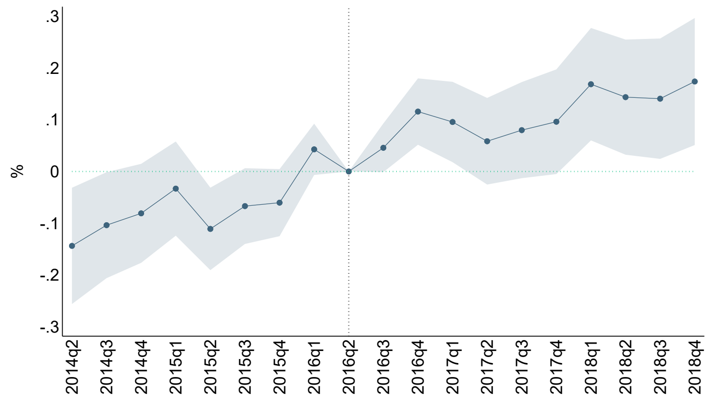


Figure 2.7. Employment.

Note: We report 95% confidence intervals. Standard errors are clustered at the firm level. The plot shows the estimated coefficients on firm exposure. Point estimates measure the percent increase in employment for every 1 p.p. increase in firm exposure. Employment is defined as the total number of formal employees in each quarter t.

2.5 A Model of Multi-Currency Bank Credit

This section presents a small open economy general equilibrium model to study the aggregate effects of large-scale foreign-currency funding shocks, such as the one induced by the tax amnesty. The framework builds on the bank lending channel model of Herreño (2023) along two dimensions. First, we extend the model to a small open economy in which firms combine labor with imported intermediate inputs in production. Second, we introduce multi-currency credit markets, allowing banks to intermediate both peso- and dollar-denominated funding and firms to borrow in either currency. The economy consists of a representative household, a finite number of non-tradable producers, a tradable sector, and a finite number of imperfectly competitive banks. As in Herreño (2023), firms must finance production expenditures in advance through working-capital loans, making production decisions sensitive to credit conditions.

A central feature of the model is that bank credit is denominated in multiple currencies, with currency-specific lending rates that depend on banks' funding conditions. Firms combine peso- and dollar-denominated borrowing to finance different production inputs, and the reliance on each currency may vary across inputs. This structure generates a direct link between the currency composition of bank funding and firms' effective marginal costs, providing a mechanism through which shifts in dollar liquidity affect the relative cost of different inputs.

We model the tax amnesty as a deposit shock that increases dollar liquidity heterogeneously across banks. Banks receiving larger inflows rely less on costly wholesale funding and reduce their dollar lending rates. Because imported inputs rely more heavily on dollar financing than wages, the decline in borrowing costs disproportionately lowers the effective cost of imports, inducing firms to expand imports and production. The resulting increase in output raises overall financing needs, generating a scale effect that also increases peso-denominated borrowing despite potentially unchanged peso lending rates. These mechanisms jointly imply differential responses in input demand and credit composition across firms, linking bank-level deposit shocks to the patterns documented in the data.

Next, we outline the model in full detail.

2.5.1 Household

The representative household consumes a final good c and supplies a total of h_{NT} hours of labor to firms in the non-tradable sector and h_T hours to the tradable sector. The household owns all firms and banks in the economy and receives total profits Π. Preferences are represented by a Greenwood-Hercowitz-Huffman (GHH) utility function:

$$U(c, h_{NT}, h_T) = \frac{1}{1-\upsilon}\left(c - \frac{h_{NT}^{1+\psi}}{1+\psi} - \Psi\frac{h_T^{1+\psi}}{1+\psi}\right)^{1-\upsilon}, \tag{2.13}$$

where υ denotes the coefficient of relative risk aversion, ψ is the inverse of the Frisch elasticity of labor supply and Ψ is a preference shifter calibrated to match the observed allocation of labor between the tradable and non-tradable sectors in the initial steady state.

The final consumption good c aggregates non-tradable consumption c_{NT} and tradable consumption c_T via a CES technology:

$$c = \left[(\Lambda_{NT})^{\frac{1}{\gamma}} c_{NT}^{\frac{\gamma-1}{\gamma}} + (1 - \Lambda_{NT})^{\frac{1}{\gamma}} c_T^{\frac{\gamma-1}{\gamma}} \right]^{\frac{\gamma}{\gamma-1}}, \tag{2.14}$$

where γ is the elasticity of substitution between tradable and non-tradable goods. The non-tradable bundle aggregates differentiated varieties produced by domestic firms:

$$c_{NT} = \left[\sum_{i=1}^{I} \Lambda_i^{\frac{1}{\sigma}} c_{i,NT}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}. \tag{2.15}$$

Standard CES demands yield the following demand schedules for non-tradable (tradable) consumption:

$$c_{NT} = \Lambda_{NT} \left(\frac{p_{NT}}{p} \right)^{-\gamma} c, \qquad c_T = (1 - \Lambda_{NT}) \left(\frac{p_T}{p} \right)^{-\gamma} c, \tag{2.16}$$

where p is the corresponding CES price index. The demand schedule for each individual variety of the non-tradable good $c_{i,NT}$ is given by:

$$c_{i,NT} = \Lambda_i \left(\frac{p_i}{p_{NT}} \right)^{-\sigma} c_{NT} \quad \forall i \in I. \tag{2.17}$$

Households supply labor to each sector according to:

$$h_{NT}^{\psi} = \frac{w_{NT}}{p}, \qquad h_T^{\psi} = \frac{w_T}{\Psi p}, \tag{2.18}$$

where w_{NT} and w_T are the prevailing wages in the non-tradable and tradable sector, respectively.

2.5.2 Firms

Non-Tradable Sector

The non-tradable sector consists of a discrete number of firms, indexed by $i \in \{1, \ldots, I\}$, producing differentiated varieties under monopolistic competition. Each firm produces output y_i

using labor h_i and an imported intermediate input M_i using a Cobb-Douglas technology:

$$y_i = A\, h_i^{\alpha_{NT}}\, M_i^{1-\alpha_{NT}}, \tag{2.19}$$

where A is aggregate productivity and $\alpha_{NT} \in (0,1)$ is the output elasticity of labor. Demand for each variety is $y_i = \Lambda_i\, (p_i/p_{NT})^{-\sigma}\, y_{NT}$.

Firms must borrow to finance input expenditures in advance. The working capital requirement, which applies to both the wage bill and the import bill, satisfies:

$$F_{i,W} = w_{NT} h_i, \qquad\qquad F_{i,M} = p_M M_i. \tag{2.20}$$

The working capital requirement for each input carries its own composite financing cost, $R_{i,W}$ and $R_{i,M}$ respectively. A key innovation of our framework is that firms optimally choose both the bank and the currency of borrowing for each input — and crucially, the currency composition of financing needs may differ across inputs. These choices are central determinants of the effective interest rate, $R_{i,W}$ and $R_{i,M}$, faced by the firm for each input which we characterize in detail in Section 2.5.3. The effective cost of employing each factor, inclusive of financing, is:

$$\widetilde{w}_i = w_{NT}(1+\theta R_{i,W}), \qquad\qquad \widetilde{p}_{i,M} = p_M(1+\theta R_{i,M}). \tag{2.21}$$

The firm maximizes profits $\Pi_i = p_i y_i - \widetilde{w}_i h_i - \widetilde{p}_{i,M} M_i$ subject to the production technology and demand. Cost minimization for a given y_i yields the unit cost function:

$$mc_{i,NT} = \frac{1}{A}\left(\frac{\widetilde{w}_i}{\alpha_{NT}}\right)^{\alpha_{NT}}\left(\frac{\widetilde{p}_{i,M}}{1-\alpha_{NT}}\right)^{1-\alpha_{NT}}. \tag{2.22}$$

Under monopolistic competition, the optimal price is a constant markup μ over marginal cost:

$$p_i = \mu \cdot \frac{1}{A}\left(\frac{w_{NT}(1+R_{i,W})}{\alpha_{NT}}\right)^{\alpha_{NT}}\left(\frac{p_M(1+R_{i,M})}{1-\alpha_{NT}}\right)^{1-\alpha_{NT}}, \tag{2.23}$$

and the demand for each factor is.:

$$h_i = \frac{\alpha_{NT}}{w_{NT}(1+R_{i,W})}\frac{1}{\mu}\, p_i y_i, \qquad\qquad M_i = \frac{(1-\alpha_{NT})}{p_M(1+R_{i,M})}\frac{1}{\mu}\, p_i y_i. \tag{2.24}$$

Financing costs act as a wedge on effective input prices, inflating the cost of each input by the input specific borrowing cost. A firm facing higher financing costs therefore demands fewer inputs, linking financial conditions to production decisions. Importantly, since the financing composite for imported inputs may differ from that of labor—reflecting the firm's currency and bank choices—financing distortions need not affect both inputs symmetrically, potentially affecting the input mix.

Tradable sector

The tradable sector consists of a representative, perfectly competitive firm that takes the world price p_T as given. Production follows a Cobb–Douglas technology:

$$y_T = A\, h_T^{\alpha_T}\, M_T^{1-\alpha_T}. \tag{2.25}$$

As in the non-tradable sector, the firm must finance its wage bill and imported inputs in advance, facing composite borrowing costs $R_{T,W}$ and $R_{X,M}$, respectively. Under perfect competition, price equals marginal cost, implying the following factor demand schedules:

$$h_T = \frac{\alpha_T}{w_T(1+R_{X,W})}\, p_T y_T, \qquad\qquad M_T = \frac{(1-\alpha_T)}{p_M(1+R_{X,M})}\, p_T y_T. \tag{2.26}$$

Market clearing in the tradable sector requires:

$$y_T = c_T + X, \tag{2.27}$$

where c_T denotes domestic demand, determined by household consumption choices in (2.16), and X denotes exports. We assume foreign demand is perfectly elastic at the world price p_T.

2.5.3 Financing Structure

Firms must finance their expenditure on each input prior to production. For each input, financing is organized in three steps. First, firms choose how to split total financing between bank credit and other external sources. This decision follows a CES structure and is common across inputs. Second, conditional on using bank credit, firms choose how to finance each input using peso- and dollar-denominated loans. The elasticity of substitution at this stage governs how easily a given input can be financed using one currency rather than the other. We allow this elasticity to differ across inputs, capturing that some expenditures are more tightly linked to foreign-currency financing. Third, for each currency, firms allocate borrowing across banks. Loans from different banks are imperfect substitutes and are combined through a CES aggregator. This allocation depends only on the currency-specific borrowing rate offered by each bank and does not vary across inputs.

Taken together, the financing structure is a three-layer CES nest: firms choose between bank and non-bank finance at the top level, determine the currency composition of bank borrowing at the intermediate level, and allocate borrowing across banks within each currency at the bottom level.

Bank Credit versus Other Finance

For each input $k \in \{W, M\}$, total financing is a CES bundle of bank loans $L_{i,k}$ and other non-bank related direct finance $OF_{i,k}$:

$$F_{i,k} = \left[\omega_B^{\frac{1}{\delta}} L_{i,k}^{\frac{\delta-1}{\delta}} + (1-\omega_B)^{\frac{1}{\delta}} OF_{i,k}^{\frac{\delta-1}{\delta}}\right]^{\frac{\delta}{\delta-1}}, \tag{2.28}$$

where ω_B is the preference for bank credit and δ is the elasticity of substitution between bank credit and other finance. For notational simplicity, we assume that ω_B does not differ across the input purposes of financing.

The composite interest rate for each k is the CES price index

$$1 + R_{i,k} = \left[\omega_B \left(1 + r_{i,L,k}\right)^{-\delta} + (1 - \omega_B)\left(1 + r_{i,OF}\right)^{-\delta}\right]^{-\frac{1}{\delta}}, \tag{2.29}$$

where $r_{i,L,k}$ is the composite bank lending rate for input k, that depends on interest rates of the banks the firm is connected to, and $r_{i,OF}$ is the rate on other external finance, assumed exogenous. The share of financing sourced from bank credit for input k is

$$S_{i,L,k} = \omega_B \left(\frac{1 + r_{i,L,k}}{1 + R_{i,k}}\right)^{-\delta}, \tag{2.30}$$

and the bank loan demand for input k satisfies $L_{i,k} = S_{i,L,k} F_{i,k}$. The cheaper bank credit is relative to alternative financing, the more the firm tilts its financing structure toward bank borrowing.

Peso versus Dollar Bank Credit

Conditional on using bank credit, firms choose the currency composition of borrowing separately for each input $k \in \{W, M\}$. Bank credit for input k is a CES aggregate of dollar- and peso-denominated loans:

$$L_{i,k} = \left[\omega_k^{\frac{1}{\kappa_k}} \left(L_{i,k}^{USD}\right)^{\frac{\kappa_k - 1}{\kappa_k}} + (1 - \omega_k)^{\frac{1}{\kappa_k}} \left(L_{i,k}^{PES}\right)^{\frac{\kappa_k - 1}{\kappa_k}}\right]^{\frac{\kappa_k}{\kappa_k - 1}}. \tag{2.31}$$

The parameter ω_k governs the baseline dollar intensity of financing input k, while κ_k governs the elasticity of substitution between dollar and peso credit for that input.

A higher ω_k means that, holding relative borrowing costs fixed, dollar credit is more useful for financing input k. By contrast, κ_k determines how easily firms can substitute between peso and dollar credit when their relative costs change. Low substitutability means that the two currencies are not merely alternative units of account, but provide different financing services. This is particularly relevant for imported inputs. Dollar credit may be better suited to finance imports because foreign suppliers often invoice, settle, or require guarantees in dollars. In that case, peso and dollar loans are imperfect substitutes even after accounting for interest-rate differences: a peso loan may not provide the same liquidity, collateral, or payment services as a dollar credit line. By contrast, wage payments are made domestically and in pesos, making peso and dollar credit more interchangeable for financing the wage bill, apart from exchange-rate risk and regulatory constraints. The model captures these differences through input-specific parameters.

The weights ω_k allow imports to have a higher baseline reliance on dollar credit than wages. The elasticities κ_k allow the ease of substitution between dollar and peso borrowing to differ across inputs. Thus, the framework can accommodate both a higher average dollar share in import financing and a heterogeneous ability to replace dollar credit with peso credit when financing different aspects of production.

The composite bank lending rate for input k is

$$1 + r_{i,L,k} = \left[\omega_k \left(1 + r_i^{USD}\right)^{1-\kappa_k} + (1 - \omega_k)\left(1 + r_i^{PES}\right)^{1-\kappa_k}\right]^{\frac{1}{1-\kappa_k}}, \tag{2.32}$$

where r_i^{USD} and r_i^{PES} are the composite interest rates that firm i faces for dollar and peso loans, respectively. These rates depend on the portfolio of banks from which the firm sources credit in each currency. The corresponding currency-specific loan demands are

$$L_{i,k}^{USD} = \omega_k \left(\frac{1 + r_i^{USD}}{1 + r_{i,L,k}}\right)^{-\kappa_k} L_{i,k}, \qquad L_{i,k}^{PES} = (1 - \omega_k)\left(\frac{1 + r_i^{PES}}{1 + r_{i,L,k}}\right)^{-\kappa_k} L_{i,k}. \tag{2.33}$$

Aggregating across inputs, total firm-level borrowing in each currency is

$$L_i^{USD} = L_i^{W,USD} + L_i^{M,USD}, \qquad L_i^{PES} = L_i^{W,PES} + L_i^{M,PES}. \tag{2.34}$$

Allocation across Banks

For each currency $\ell \in \{USD, PES\}$, firm i's total demand L_i^ℓ is sourced from banks via a competitive CES aggregator:

$$L_i^\ell = \left[\sum_b s_{ib}^{\frac{1}{\varepsilon}} \left(L_{ib}^\ell \right)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}. \tag{2.35}$$

The parameter ε governs the elasticity of substitution across loans from different banks. The weights $s_{ib} \geq 0$, with $\sum_b s_{ib} = 1$, represent the strength of the lending relationship between firm i and bank b, assumed common across currencies. The CES structure captures imperfect substitutability across bank loans due to frictions in switching and relationship-based intermediation.

The effective interest rate firm i faces in currency ℓ is

$$1 + r_i^\ell = \left[\sum_b s_{ib} \left(1 + r_b^\ell \right)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}, \tag{2.36}$$

and the demand for loans from bank b in currency ℓ satisfies:

$$L_{ib}^\ell = s_{ib} \left(\frac{1 + r_b^\ell}{1 + r_i^\ell} \right)^{-\varepsilon} L_i^\ell \qquad \forall i, b, \ell. \tag{2.37}$$

2.5.4 Banking Sector

Banks are indexed by $b \in \{1, \ldots, B\}$ and operate under monopolistic competition in the loan market. Each bank sets currency-specific lending rates, taking as given the demand schedules generated by firm-level loan aggregators. Banks maintain separate balance sheets for peso and dollar operations:

$$L_b^{PES} = D_b^{PES} + N_b^{PES}, \tag{2.38}$$

$$L_b^{USD} = D_b^{USD} + K_b + N_b^{USD}, \tag{2.39}$$

where D_b^ℓ denotes deposits, N_b^ℓ denotes wholesale funding in currency ℓ, and K_b is bank equity, which we allocate to the dollar book. Dollar deposits include the amnesty-induced inflow:

$$\Delta D_b^{USD} = D_{b,post}^{USD} - D_{b,pre}^{USD} = \text{Amnesty Funds}_b. \tag{2.40}$$

To capture how bank funding composition affects lending behavior, we assume that wholesale funding costs decline with deposit liquidity. This is consistent with evidence that wholesale funding becomes more expensive as the deposit base erodes. The marginal cost of funds is currency specific:

$$MC_b^{PES} = 1 + r^{*,PES} - \phi_{PES}\frac{D_b^{pesos}}{B_b^{PES}} + \rho, \tag{2.41}$$

$$MC_b^{USD} = 1 + r^{*,USD} - \phi_{USD}\frac{D_b^{USD}}{B_b^{USD}} + \rho, \tag{2.42}$$

where $r^{*,\ell}$ are exogenous base rates, ρ is a risk premium, ϕ_{USD} governs the sensitivity of dollar funding costs to dollar deposit scarcity, ϕ_{PES} governs the sensitivity of peso funding costs to peso deposit scarcity, and B_b^ℓ denotes a currency-specific benchmark liquidity level. The currency-specific benchmarks allow the model to capture that peso and dollar wholesale funding markets may evaluate bank liquidity differently. Peso wholesale funding is sourced domestically, through interbank markets, central bank facilities, and local institutional investors, so its cost depends on the bank's peso liquidity position. Dollar wholesale funding is instead obtained from international counterparties, such as correspondent banks, which potentially place greater weight on the bank's foreign-currency liquidity. To align the model with our empirical exposure measure, we set the dollar benchmark to pre-amnesty total deposits, $B_b^{USD} = D_{b,pre}^{Total}$. Under this normalization,

$$\Delta MC_b^{USD} = -\phi_{USD} \frac{\Delta D_b^{USD}}{D_{b,pre}^{Total}} = -\phi_{USD} S_b. \tag{2.43}$$

Thus, dollar deposit inflows lower dollar funding costs in direct proportion to the empirical exposure S_b, while changes in peso deposits do not mechanically affect dollar rates. This specification should be interpreted as a reduced-form normalization that preserves the mapping between the model and the identifying variation in the data. Monopolistic competition in the lending market implies that optimal loan rates are set as a constant markup over marginal cost:

$$1 + r_b^{PES} = \frac{\varepsilon}{\varepsilon - 1} \left[1 + r^{*,PES} - \phi_{PES} \frac{D_b^{PES}}{B_b^{PES}} + \rho \right], \tag{2.44}$$

$$1 + r_b^{USD} = \frac{\varepsilon}{\varepsilon - 1} \left[1 + r^{*,USD} - \phi_{USD} \frac{D_b^{USD}}{B_b^{USD}} + \rho \right]. \tag{2.45}$$

These expressions imply cross-bank heterogeneity in lending rates, driven by differences in deposit liquidity across currencies. Banks with greater deposit funding in a given currency face lower marginal costs and therefore offer more favorable lending rates in that market segment. Aggregating across borrowers, total credit demand faced by bank b in each currency is given by:

$$L_b^{USD} = \sum_i L_{ib}^{USD}, \qquad L_b^{PES} = \sum_i L_{ib}^{PES}. \tag{2.46}$$

Total lending by bank b is then

$$L_b = L_b^{USD} + L_b^{PES}. \tag{2.47}$$

2.5.5 Market Clearing and Aggregation

We close the model by imposing market clearing conditions and defining aggregate variables. Labor markets clear in each sector:

$$h_{NT} = \sum_{i=1}^{I} h_i, \qquad\qquad h = h_{NT} + h_T. \tag{2.48}$$

Non-tradable goods markets clear variety by variety, $c_{i,NT} = y_i$ for all i, which implies $c_{NT} = y_{NT}$ in aggregate.

Total imports equal the sum of imported inputs used by all firms:

$$\mathcal{M} = \sum_{i=1}^{I} M_i + M^X. \tag{2.49}$$

The balance of payments equates net exports to net financial outflows:

$$p_T(y_T - c_T) - p_M\mathcal{M} = \sum_b \left[r^{*,USD} - \phi_{USD}\frac{D_b^{USD}}{B_b} \right] N_b^{USD}, \tag{2.50}$$

where the right-hand side captures interest payments on dollar-denominated wholesale funding.

Nominal GDP is given by the sum of domestic absorption and the trade balance:

$$pGDP = p\,c + p_T(y_T - c_T) - p_M\mathcal{M}. \tag{2.51}$$

Normalizing $p_T = p_M = 1$, real GDP in units of the final consumption bundle is

$$GDP = c + \frac{(y_T - c_T) - \mathcal{M}}{p}. \tag{2.52}$$

With tradable prices normalized to one, the real exchange rate is given by

$$rer = \frac{p_T}{p} = \frac{1}{p}. \tag{2.53}$$

Using this definition, real GDP can be expressed as

$$GDP = c + rer\left[(y_T - c_T) - \mathcal{M}\right]. \tag{2.54}$$

2.5.6 Equilibrium

Given the deposit shock $\{\text{Amnesty}_b\}_b$, the external financing rate $r_{i,OF}$, bank equity $\{K_b\}_b$, and parameters, an equilibrium consists of prices $\{p_i, p_{NT}, p_T, p, w_{NT}, w_T\}$, interest rates $\{r_b^{USD}, r_b^{PES}, r_i^{USD}, r_i^{PES}, r_{i,L,k}, R_{i,k}\}$, and real allocations $\{y_i, h_i, M_i, c_{i,NT}, c_T, c, h, y_T, h_T, M_T, X, L_{ib}^{\ell}, L_{i,k}^{\ell}, N_b^{\ell}\}$ such that all agents optimize and all markets clear.

2.5.7 Mapping Empirical Estimates to Structural Objects

We use the model to derive structural expressions for the reduced-form coefficients estimated in Sections 2.4.2 and 2.4.3. To simplify notation, we show derivations assuming an initial equilibrium with symmetric rates ($r_b^{USD} = r_b^{PES} = r_{i,OF} = \bar{r}$ for all b and all i), however, in the quantitative exercise we allow for a wedge between peso and dollar rates, as well as heterogeneous baseline bank rates, consistent with the context that we study. Throughout this section, we drop the distinction between tradable and non-tradable firms and we denote the production-weighted average dollar preference by

$$\bar{\omega} \equiv \alpha\,\omega_W + (1-\alpha)\,\omega_M \tag{2.55}$$

where ω_W and ω_M are the dollar preferences in the wage and import bank financing aggregators, respectively, and α denotes the labor share. As mentioned earlier, we set the benchmark level of deposits B_b to the pre-amnesty level of total deposits $B_b = D_{b,pre}^{Total}$. For small r, this implies that the change in the bank-level dollar rate between the pre- and post amnesty shock satisfies:

$$
\begin{aligned}
d\ln(1+r_b^{USD}) \approx dr_b^{USD} &= -\phi_{USD}\frac{\Delta D_b^{USD}}{D_0^{Total}} \\
&= -\phi_{USD} \times S_b,
\end{aligned}
\tag{2.56}
$$

where S_b matches the measure of bank-level exposure employed in Section 2.4.

Bank Lending Channel

We log-linearize the bank–firm loan allocation equation to isolate the effect of dollar deposit shocks on credit supply within the dollar segment of the credit market. The model-analogue for the bank lending channel regression is:

$$
d\ln L_{ib}^{USD} = \varepsilon\,\phi_{USD}\,S_b + \lambda_i + e_{ib}. \tag{2.57}
$$

This expression captures how bank-level funding shocks translate to within-firm credit reallocation across lenders. Conditional on firm-level credit demand and aggregate financing conditions, banks receiving larger dollar deposit inflows reduce their lending rates and expand dollar credit supply to a given firm. The strength of this reallocation depends on two key parameters: the elasticity of substitution across banks, ε, and the sensitivity of dollar lending rates to bank liquidity, ϕ_{USD}.

To connect this expression to our empirical specification, we exploit the assumption that the change in peso lending rates is common across banks. As a result, changes in equilibrium peso lending to a given firm are spread homogeneously across banks,

$$
d\ln L_{ib}^{PES} = d\ln L_i^{PES}, \tag{2.58}
$$

and are absorbed by firm fixed effects. Identification therefore comes entirely from variation in dollar lending rates across banks within a firm. Because our data do not separately measure lending by currency at the bank–firm level, we work with total lending $L_{ib} = L_{ib}^{USD} + L_{ib}^{PES}$. A first-order approximation around the initial equilibrium yields:

$$d \ln L_{ib} = \bar{\omega}\, \varepsilon \phi_{USD}\, S_b + \lambda_i + e_{ib}, \tag{2.59}$$

where $\bar{\omega}$ denotes the baseline share of dollar lending in total bank credit. This rescaling relative to expression (2.57) reflects that the deposit shock operates only through the dollar segment of lending, so its effect on total credit is proportional to the initial weight of dollar-denominated loans. Equation (2.59) provides a direct mapping to our empirical bank lending channel regression:

$$\beta^{BLC} = \bar{\omega}\, \varepsilon\, \phi_{USD}. \tag{2.60}$$

Firm Borrowing Channel

The response of total bank credit at the firm level reflects two distinct margins: (i) an expansion in overall financing needs driven by higher production, and (ii) a reallocation from other external finance toward bank credit as the cost of bank credit declines. The model-implied analogue of the firm-borrowing regression is:

$$d \log L_i = \lambda + \phi_{USD}\, \bar{\omega}\, \left[\sigma\, \omega_B + \delta\, (1 - \omega_B)\right] S_i + e_i, \tag{2.61}$$

where λ is an intercept that captures general-equilibrium effects common across firms.

The implied elasticity is therefore:

$$\beta^{FB} = \phi_{USD}\, \bar{\omega}\, \left[\sigma\, \omega_B + \delta\, (1 - \omega_B)\right]. \tag{2.62}$$

The first component, $\sigma\, \omega_B$, captures the scale effect: lower borrowing costs reduce marginal costs and prices, increasing demand and, in turn, financing needs. This channel is stronger when goods are more substitutable, as a given price decline translates into a larger expansion in firm-level output. The second component, $\delta(1-\omega_B)$, captures substitution across financing sources. As bank lending rates fall relative to other external finance, firms reallocate borrowing toward banks. This effect is stronger when bank credit and other external finance are close substitutes (i.e. higher δ). Both margins are scaled by the baseline dollar exposure $\bar{\omega}$, reflecting that the deposit shock operates through the dollar component of bank lending. As a result, the pass-through from bank-level funding shocks to firm-level borrowing is proportional to the initial importance of dollar-denominated credit in firms' financing structure.

Currency Composition Channel

We next characterize how firms adjust the currency composition of their borrowing in response to the shock. The change in the share of dollar credit can be decomposed into two distinct channels:

$$d\ln\left(\frac{L_i^{usd}}{L_i}\right) = \lambda - \sum_k \underbrace{\theta_k^{usd}\kappa_k(1-\omega_k)\,d\ln(1+r_i^{usd})}_{\text{Within-input substitution}} + \sum_k \underbrace{[\theta_k^{usd}-\alpha_k]\,d\ln L_i^k}_{\text{Differential FBC across inputs}} + e_i \quad (2.63)$$

where α_k denotes the output share of input k (i.e. α and $(1-\alpha)$). Here, $\theta_k^{usd} = \frac{\omega_k \alpha_k}{\bar{\omega}}$ denotes the share of input k in total dollar borrowing, so that the term $(\theta_k^{usd} - \alpha_k)$ captures whether input k is more or less dollar-intensive than predicted by its participation in production.

The first term captures within-input substitution: as dollar borrowing becomes cheaper, firms tilt the financing of each input toward dollar-denominated credit. The strength of this channel is governed by the elasticity of substitution across currencies, κ_k, and the baseline reliance on peso financing, $1-\omega_k$. The second term captures across-input reallocation. When inputs differ in their dollar intensity, changes in the composition of production and financing needs affect the overall currency mix of borrowing. In particular, if inputs that are more dollar-intensive expand more in response to the shock, the aggregate share of dollar borrowing increases even absent within-input substitution.

Substituting for $d\ln L_i^k$ yields the following estimating equation:

$$d\ln\left(\frac{L_i^{usd}}{L_i}\right) = \lambda + \phi_{USD}\left[\sum_k \theta_k^{usd}\kappa_k(1-\omega_k) + \sum_k\left(\theta_k^{usd}-\alpha_k\right)\omega_k\,\delta\right]S_i + e_i \tag{2.64}$$

where λ captures general-equilibrium effects common across firms. The implied elasticity is therefore:

$$\beta^{CC} = \phi_{USD}\left[\sum_k \theta_k^{usd}\kappa_k(1-\omega_k) + \sum_k\left(\theta_k^{usd}-\alpha_k\right)\omega_k\,\delta\right] \tag{2.65}$$

When inputs have identical dollar intensities ($\omega_k = \omega$), the across-input channel is absent. If substitution across currencies is also homogeneous ($\kappa_k = \kappa$), the elasticity reduces to:

$$\beta^{CC} = \phi_{USD}\,\kappa. \tag{2.66}$$

Intuitively, the share of dollar credit responds more when it is easier to substitute bank credit denominated in different currencies.

Input Demand

We next characterize the response of firm-level input demand to the deposit shock. The responses of employment and imported inputs combine two channels. First, the shock lowers average financing costs, reduces marginal costs, and expands output. This generates a scale effect that raises demand for both inputs. This is the same scale effect analyzed for the firm borrowing channel. Second, when imports and wages differ in their dollar financing intensity, the shock changes their relative effective costs which generates a substitution effect. If imported inputs are more dollar-intensive than wage payments, the decline in dollar borrowing rates lowers the effective cost of imported inputs by more than the effective cost of labor, inducing substitution toward imports. The model-implied elasticities for each input are:

$$\beta^{Emp} = \underbrace{\phi_{USD}\omega_B\sigma\bar{\omega}}_{\text{scale effect}} + \underbrace{\phi_{USD}\omega_B(1-\alpha)(\omega_W - \omega_M)}_{\text{substitution effect}}, \tag{2.67}$$

$$\beta^{Imp} = \underbrace{\phi_{USD}\omega_B\sigma\bar{\omega}}_{\text{scale effect}} + \underbrace{\phi_{USD}\omega_B\alpha(\omega_M - \omega_W)}_{\text{substitution effect}}, \tag{2.68}$$

The scale effect is common across inputs and increasing in the elasiticity of substitution σ, as discussed for the firm borrowing channel. The substitution effects have opposite signs across inputs. When $\omega_M > \omega_W$, imported inputs benefit more from the decline in dollar borrowing costs, so firms substitute toward imports and away from labor, holding the price of each input fixed. The difference between the import and employment elasticities isolates this relative-cost channel:

$$\beta^{Imp} - \beta^{Emp} = \phi_{USD}\omega_B(\omega_M - \omega_W). \tag{2.69}$$

This gap is strictly positive if import financing is more dollar-intensive than wage financing. These expressions are derived around the symmetric benchmark. Outside this benchmark, the elasticities also depend on the initial peso–dollar interest rate wedge and on the elasticity of substitution between peso and dollar credit for each input.[14]

[14] Outside of the symmetric equilibrium, the expression for these elasticities is given by:

$$\beta^W = -\phi^{usd}\omega_B \times \left[\frac{\Lambda_W[(1-\sigma)\alpha - 1]\omega_W}{\omega_W + (1-\omega_W)\Delta_W^{usd}} + \frac{\Lambda_M(1-\sigma)(1-\alpha)\omega_M}{\omega_M + (1-\omega_M)\Delta_M^{usd}}\right]$$

$$\beta^M = -\phi^{usd}\omega_B \times \left[\frac{\Lambda_M[(1-\sigma)(1-\alpha) - 1]\omega_M}{\omega_M + (1-\omega_M)\Delta_M^{usd}} + \frac{\Lambda_W(1-\sigma)\alpha\omega_W}{\omega_W + (1-\omega_W)\Delta_W^{usd}}\right]$$

with $\Delta_k^{usd} = \left(\frac{1+\bar{r}^{PES}}{1+\bar{r}^{USD}}\right)^{1-\kappa_k}$, $\Lambda_M = \left(\frac{1+\bar{r}_i^{L,M}}{1+\bar{r}_i^{M}}\right)^{1-\delta}$ and $\Lambda_W = \left(\frac{1+\bar{r}_i^{L,W}}{1+\bar{r}_i^{W}}\right)^{1-\delta}$. Throughout, a bar over an interest rate denotes its value in the initial equilibrium, before the deposit shock.

Identifying structural parameters

Collecting the previous elasticities, the model implies that:

$$\beta^{BLC} = \phi_{USD}\,\bar{\omega}\,\varepsilon \tag{2.70}$$

$$\beta^{FB} = \phi_{USD}\,\bar{\omega}\,[\sigma\,\omega_B + \delta\,(1-\omega_B)] \tag{2.71}$$

$$\beta^{CC} = \phi_{USD}\left[\sum_k \theta_k^{usd}\kappa_k(1-\omega_k) + \sum_k \left(\theta_k^{usd} - \alpha_k\right)\omega_k\,\delta\right] \tag{2.72}$$

$$\beta^{Emp} = \phi_{USD}\,\omega_B\,[\omega_W + (\sigma-1)\bar{\omega}] \tag{2.73}$$

$$\beta^{Imp} = \phi_{USD}\,\omega_B\,[\omega_M + (\sigma-1)\bar{\omega}] \tag{2.74}$$

Since ϕ_{USD} enters all moments multiplicatively, ratios of estimated responses identify the substitution parameters ε, δ, κ_W, and κ_M, while the absolute scale of any single moment recovers ϕ_{USD}. Outside the symmetric equilibrium, the employment and import elasticities also depend on κ_W and κ_M through the initial peso–dollar interest rate wedge, providing the additional variation needed to identify the currency-substitution elasticities given δ recovered from β^{FB}. This gives a system of five moments in five parameters. In practice, identification of κ_M is weak because import financing in Argentina is heavily concentrated in dollars, leaving little currency-composition variation on that margin. We therefore fix $\kappa_M = 1$ and estimate $(\phi_{USD}, \varepsilon, \delta, \kappa_W)$ by minimum distance targeting all five moments, with β^{CC} providing an over-identifying restriction.

2.5.8 Calibration

We calibrate the model to reflect the structural and financial characteristics of the Argentine economy using a combination of administrative data and aggregate national accounts. Parameters are grouped into four categories: non-bank parameters, banking sector parameters, loan allocation weights, and steady-state ratios. Table 2.8 reports the values and sources. All variables are expressed in real terms and reflect pre-amnesty averages.

We set σ, the elasticity of substitution across firm varieties, to 7 which implies a 15% steady-state markup. For the Frisch elasticity, $\frac{1}{\psi}$, we consider three alternative values to assess the sensitivity of our results to labor supply responses.

We calibrate production elasticities using the observed input cost shares. With a Cobb–Douglas technology, output elasticities correspond to cost shares evaluated at marginal cost; in our setting, this would ideally include financing costs. In practice, we approximate these elasticities using expenditure shares on labor and imported inputs, abstracting from financing costs due to the lack of input-specific interest rate data.[15] Specifically, we group non-importing firms and importing firms that do not export into the non-tradable sector and compute the aggregate share of the wage bill in total input expenditures to discipline α_{NT}. We then apply the same procedure to exporting firms to calibrate α_T.[16]

We calibrate the parameters that govern the composition of firm financing as follows. We use firms that do not import or export to discipline wage-financing parameters. For these firms, bank borrowing primarily reflects wage financing, so the ratio of credit to the wage bill identifies the bank-financing intensity for labor, $\omega_{B,W}$, while their dollar share of borrowing identifies the currency composition of wage financing, ω_W. We then turn to importing firms, for which total credit finances both wages and imported inputs. Using data on their total credit, wage bill, and imports, and imposing that the wage-financing intensity is the same as for non-importers, we back out the bank-financing intensity of imports, $\omega_{B,M}$. Finally, combining these objects with the observed dollar share of borrowing among importers, we recover the dollar share of import financing, ω_M, as the residual that reconciles the aggregate currency composition of credit with its decomposition into wage and import financing. This procedure relies on the assumption

[15]By omitting domestic intermediates and calibrating using only labor and imported inputs, we effectively normalize the cost shares over a reduced set of inputs. This implies that the levels of both elasticities are inflated relative to their true values, as the denominator excludes the share of domestic intermediates. However, the relative importance of labor versus imported inputs is preserved, with a Cobb-Douglas technology.

[16]Using raw expenditure shares instead of financing-inclusive cost shares introduces a potential bias. If the effective financing wedge is higher for wages than for imports, this procedure understates the labor elasticity α_s, while the opposite holds if the effective financing wedge is higher for imports. Because imports may face lower interest rates but a larger fraction of expenditures requiring financing, the direction of the bias is theoretically ambiguous.

that differences in credit intensity across firms arise from input composition rather than from systematic differences in financing needs across firm types.

We solve the model for $B = 4$ and $I = 100$. We calibrate the exposure of each bank as follows. In the data, we rank banks by exposure and bin them into four groups, each accounting for one fourth of total lending at the start of the tax amnesty. This leaves us with four bank types. We then rank firms by exposure and bin them into 100 firm types, each accounting for one percent of total lending. For each firm bin, we compute the share of lending from each bank type at baseline and compute bin-level exposure as in the empirical section. We choose equal weights for each firm-bin in the consumption index.

The rest of the banking parameters are set as follows. We calibrate the shares of deposits, equity, and whole-sale funding using bank-balance sheet data. During the period of analysis, deposits made up 70% of aggregate bank funding and equity an additional 10%. Since we do not have precise data on foreign whole-sale funding, we calibrate it as the residual 20%. We assume equal balance-sheet composition across banks before the amnesty.

We model the tax amnesty as an exogenous increase in aggregate deposits that accrue to individual banks with heterogeneous intensity. We calibrate the size of the shock to match the aggregate increase in deposits, both as a share of GDP and as a share of bank funding, and its distribution across banks as explained above. We hold all other exogenous variables fixed at their pre-amnesty levels, this includes the baseline interest rates $r^{*,\ell}$ in each currency, bank-level equity K_b, the interest rate on external finance $r_{i,OF}$, and aggregate productivity A.

2.5.9 Aggregate Effects of the Financial Channel of the Tax Amnesty

We use the model to evaluate the aggregate effects of the credit supply shock caused by the Tax Amnesty. We simulate an exogenous increase in bank deposits, consistent with the observed inflows during Argentina's 2016 tax amnesty. The model captures how this liquidity expansion affects loan supply, interest rates, firms' currency composition, input demand, and real activity through general equilibrium adjustments.

Table 2.8. Calibration.

Symbol	Description	Value	Source
	Household Block		
$1/\psi$	Frisch elasticity of labor supply	*varies*	Sensitivity analysis
γ	EoS between tradables and non-tradables	0.5	Gonzalez-Rozada et al. (2003)
σ	EoS across varieties	7	Mark-up of 17%
Λ_i	Firm i weight in consumption index	$1/I$	Normalization
Λ_{NT}	Weight of non-tradables in consumption	0.77	National Accounts
	Production Block		
α_{NT}	Labor share in non-tradable production	0.93	Cost share data
α_T	Labor share in tradable production	0.79	Cost share data
$\omega_{B,NT}$	Share of bank credit in financing - Non-tradables	.27	Credit registry
$\omega_{B,T}$	Share of bank credit in financing - Tradables	.51	Credit registry
$\omega_{k,NT}$	USD share in k's financing - Non-tradables	.05(.91)	Credit registry (See text)
$\omega_{k,T}$	USD share in k's financing - Tradables	.05 (.95)	Credit registry (See text)
s_{ib}	Bank-firm loan shares		Credit registry
κ_k	EoS across currencies for financing input k	3.5 (1)	Estimated (See text)
ε	EoS across banks	7.5	Estimated
δ	EoS across bank and other finance	6.75	Estimated
	Bank Block		
ϕ_{USD}	Sensitivity of dollar funding costs to deposits	0.85	Estimated
S_b	Bank exposure to amnesty inflows	varies by bank	Bank balance sheet data
$r^{*,USD}$	Baseline wholesale dollar funding rate	.04	Average real dollar rate
$r^{*,PES}$	Baseline wholesale peso funding rate	.08	Average real peso rate
B_b	Benchmark liquidity level	varies by bank	Total deposits pre-shock
	Macro Aggregates		
Ω_X	Share of exports in GDP	0.12	National Accounts
Ω_M	Share of imports in GDP	0.12	National Accounts
Ω_C	Domestic absorption to GDP	1	Normalization

Note: EoS denotes elasticity of substitution. The set of parameters $\Omega^{\ell}_{b,j}, \Omega_X, \Omega_M, \Omega_C$ are required to solve the model using hat algebra.

In the baseline calibration, we set the inverse Frisch elasticity to one and estimate the structural parameters to match the reduced-form elasticities as described in Section 2.5.7. The model implies that the tax-amnesty deposit shock increased bank loans by 6.4% and GDP by .91%. This corresponds to an output-to-loans elasticity of 0.14. As a benchmark, we also report a back-of-the-envelope calculation that applies a general-equilibrium multiplier of one to the reduced-form employment and import elasticities, weighted by their initial-equilibrium cost shares, implying a GDP increase of 2.1%, larger than the model-implied effect.

Table 2.9 reports the fit of the five targeted moments at the estimated parameters. The model matches the bank lending channel, firm borrowing, currency composition, and employment elasticities to within 0.05. The largest discrepancy is on the import elasticity, which the model undershoots by 0.14. Raising κ_M to 7 (i.e. setting it equal to $\hat{\varepsilon}$) closes only part of this gap,[17] suggesting the model may be missing an extensive margin of import adoption or understating substitutability between domestic and imported inputs in production.

Table 2.9. Moment Fit.

Moment	Target	Model	Gap
Bank lending channel (β^{BLC})	1.070	1.070	−0.000
Firm borrowing (β^{FB})	0.970	1.011	+0.041
Currency composition (β^{CC})	0.600	0.577	−0.023
Employment (β^{Emp})	0.170	0.185	+0.015
Imports (β^{Imp})	1.230	1.089	−0.141

Note: The table reports the targeted reduced-form moments, the model-implied values at the estimated parameters ($\phi_{USD}, \varepsilon, \delta, \kappa_W$), and the gap between the two.

Table 2.10 reports the baseline results and a set of counterfactual simulations, organized in three groups. The first group varies the currency-substitution elasticity for wage financing, κ_W. The second group varies labor supply flexibility through the Frisch elasticity. The third varies the degree of openness. Together, these exercises characterize how the real effects of a dollar liquidity shock depend on the supply-side margins available to firms.

One key margin in the model is how easily firms can substitute between peso and dollar

[17] Setting $\kappa_M = 7$ at the baseline calibration changes GDP from 0.92% to 0.94% and raises the model-implied import elasticity from 1.09 to 1.11.

credit when financing wage payments. Because import financing is already heavily dollarized at baseline, the currency-substitution elasticity for imports κ_M has little scope to operate, and we fix it at the Cobb-Douglas limit of one. The interesting variation is therefore in κ_W, which governs how much of the decline in dollar borrowing costs is transmitted to firms' effective cost of hiring.

The baseline estimate of $\kappa_W = 3.5$ implies that firms have moderate ability to substitute toward dollar credit when financing wages, and the model implies a GDP increase of .92%. To assess how much this margin matters for aggregate transmission, we consider two alternative scenarios. Setting $\kappa_W \approx 1$ (i.e. Cobb-Douglas) which implies firms maintain fixed currency shares in wage financing and the dollar shock passes through only via the initial dollar exposure ω_W. In this scenario, GDP increases by .81%. At the other extreme, setting $\kappa_W = \varepsilon \approx 7$ makes peso-dollar substitutability for wage financing as high as substitutability across banks and results in a GDP increase of 1.12%. The range of .8 to 1.12 percentage points across these three scenarios indicates that the currency-substitution margin on the wage side is a quantitatively important determinant of aggregate transmission.

We next ask how the aggregate response varies with the supply-side elasticity of the economy, holding the financial structure fixed at the baseline. Two parameters govern this elasticity: the Frisch elasticity of labor supply and the degree of openness.

A higher Frisch elasticity allows the labor market to accommodate the expansion in firms' hiring demand with smaller wage adjustment, amplifying the quantity response. Reducing the inverse Frisch elasticity from one to two-thirds dampens the GDP response from .91% to .78%.

Openness plays an analogous role through the imports margin. When the economy is more open, firms can expand production by drawing on imported inputs supplied at fixed world prices, making the effective supply of production inputs more elastic. This is the open-economy counterpart to a high Frisch elasticity: both increase the responsiveness of output to a given reduction in financing costs. Reducing openness from $(X + M)/\text{GDP} = 0.24$ to 0.10 lowers the GDP response to .45%, as firms become more reliant on domestic labor for expansion and the tighter labor market constrains the output response.

Table 2.10. Simulation Results: Aggregate Effects of the Tax Amnesty.

	Currency Substitution			Labor Supply	Openness
	$\kappa_W = 1$	$\kappa_W \approx 3$ **(Baseline)**	$\kappa_W = \varepsilon \approx 7$	$1/\psi = 2/3$	$(X+M)/\text{GDP} = 0.10$
Back-of-the-envelope ($\mathcal{M} = 1$)		2.1%			
GDP	0.81%	0.92%	1.12%	0.78%	0.45%
Tradable output	2.15%	2.10%	1.98%	1.91%	0.99%
Non-tradable output	1.08%	1.25%	1.58%	1.13%	0.56%
Tradable/Non-tradable output	1.99	1.68	1.25	1.69	1.76
Bank loans	6.12%	6.40%	6.93%	6.24%	3.80%
Y/L elasticity	0.13	0.14	0.16	0.13	0.12

Note: The table shows percentage changes relative to the initial steady state following the tax-amnesty deposit shock.

Overall, the simulations suggest that capital inflows associated with asset repatriation can have non-negligible aggregate effects when the financial system intermediates them into credit. The magnitude of the response depends on the currency-substitution margin available to firms for wage financing, the flexibility of labor supply, and the degree to which firms can substitute toward imported inputs at world prices. The latter two operate symmetrically as supply-side elasticities that amplify or dampen the transmission of financial shocks to real activity.

2.6 Conclusion

This paper identifies a new channel through which tax amnesties affect the economy: the financial channel. When designed to encourage asset repatriation, these policies can expand the financial sector, especially in underdeveloped credit markets. We study Argentina's 2016 Tax Amnesty, one of the largest in history, and show that the inflow of hidden assets into domestic banks increased credit supply. Banks more exposed to these inflows expanded lending, which in turn improved firms' access to credit. This facilitated firm growth, as more exposed firms raised imports, exports, and employment.

The structural model we develop clarifies the mechanism behind these effects. By allowing repatriated assets to strengthen domestic bank balance sheets, tax amnesties expand the supply of credit, which then relaxes firms' access to credit. In our preferred calibration, the resulting credit expansion translates into a real GDP gain of around .9%, highlighting the macroeconomic significance of the financial channel. This mechanism provides a new rationale for implementing tax amnesty policies—beyond their direct fiscal impact—particularly in economies where credit markets remain shallow.

Taken together, our findings reveal that tax amnesty policies can stimulate economic growth by expanding the financial sector, demonstrating effects beyond their direct fiscal impact. These results are particularly relevant for countries with underdeveloped financial systems, where the potential for growth through improved access to capital is significant.

The financial channel of tax amnesties has two main policy implications. First, it highlights an additional benefit that should be considered when evaluating such programs. Second, it suggests that policymakers should design incentives not only to encourage disclosure of hidden assets but also to promote their repatriation into domestic banks.

These findings have broader implications beyond tax amnesties. They demonstrate that sudden liquidity shocks, particularly in foreign currency, can materially influence credit availability, firm growth, and aggregate economic activity. These inflows disproportionately benefit firms engaged in international trade. Policymakers should consider both the incentives for asset repatriation and the potential effects of capital inflows on the domestic financial system when evaluating macroeconomic policy tools.

Acknowledgements

Chapter 2, in full, is based on the paper from Federico Bernini, Paula Donaldson, Ezequiel Garcia-Lembergman, and Leticia Juarez, "The Financial Channel of Tax Amnesty Policies" currently being prepared for submission. The dissertation author is one of the primary authors of this paper.

Chapter 3

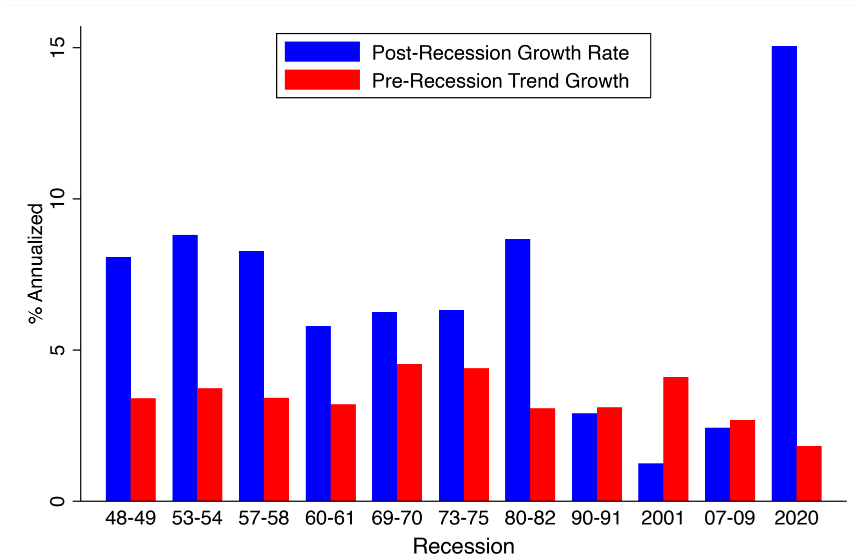
Why are Some Recoveries Weak and Others Strong?

3.1 Introduction

The recoveries from the 1990-1, 2001, and 2007-9 recessions have been weak. The blue bars in Figure 3.1 show that private domestic demand[1] grows by only 2.2% on average in the four quarters following these recessions, much less than the 8.4% average growth rate in the four quarters following other recessions. The red bars show that growth after each of the 1990-1, 2001, and 2007-9 recessions is also weak relative to the pre-recession trend, computed as the annualized growth rate from the last business cycle peak to the current business cycle peak. This contrasts with the strong recoveries from the other recessions, in which recovery growth exceeds the pre-recession trend. This paper asks why are the recoveries from the 1990-1, 2001, and 2007-9 weak—recovery growth is below the pre-recession trend—and why are recoveries from the other recessions strong?

A range of explanations have been put forward to explain why some recoveries are weak: Bad luck (Galí et al., 2012), investment boom-bust cycles (Beaudry et al., 2018; Rognlie et al., 2018), credit boom-bust cycles (Jordà et al., 2013), different sectoral incidences of shocks (Beraja and Wolf, 2022), changes in beliefs (Kozlowski et al., 2019, 2020), the secular decline in manufacturing (Leamer, 2021), hysteresis (Benigno and Fornaro, 2018; Garga and Singh, 2021;

[1] The sum of consumption expenditures and private investment.



Note: The post-recession growth rate is the growth rate of private domestic demand (consumption expenditures plus private investment) for the four quarters after the recession trough. The pre-recession trend is the annualized growth rate between the prior business cycle peak and the current business cycle peak, except for 1948-9 where the trend is linearly extrapolated from 1947Q1-1948Q4.

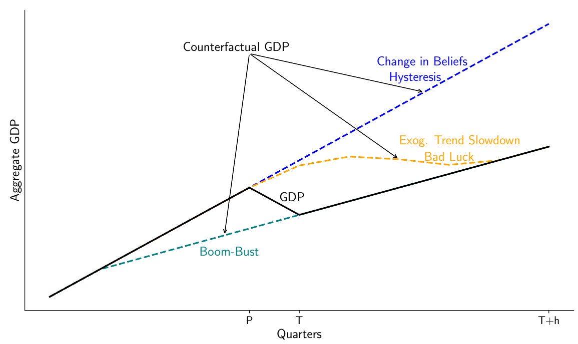
Figure 3.1. Growth of Private Domestic Demand after Recession Trough.

Reifschneider et al., 2015), and an exogenous slowdown in trend growth (Fernald et al., 2025).

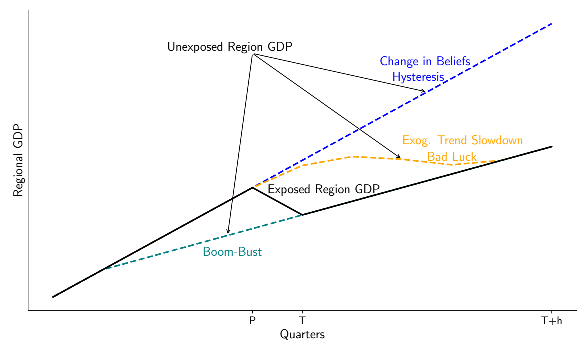
These theories imply different counterfactual paths for output. Figure 3.2A illustrates these counterfactuals around a hypothetical GDP path similar to the Great Recession. Hysteresis or changes in beliefs due to the recession imply a persistent gap between realized GDP and its counterfactual path absent the recession. An exogenous slowdown in trend growth instead implies that growth would have been weak even absent the recession. We also expect the counterfactual and realized GDP path to converge as bad luck eventually ends. And boom-bust explanations imply that GDP was high before the recession, and through the recession converges to counterfactual GDP.

Distinguishing between these theories is difficult because the counterfactual path of output is unobserved. Common strategies to recover aggregate counterfactuals include extrapolating from existing trends (e.g., Blanchard et al., 2015), extrapolating from other countries or past recessions (e.g., Jordà et al., 2013), or using a structural model to infer shocks and counterfactuals (e.g., Galí et al., 2012). That disagreement about the correct counterfactual remains reflects that each of these methods embodies important trade-offs. Trend extrapolation holds fixed the country and the recession, but must assume that the trend continues on its existing path. In practice, trend extrapolations are sensitive to the chosen horizon (Ball, 2015), and would be invalid under some theories such as boom-bust cycles or if aggregate output is non-stationary. In comparing a recession to past recessions or across countries one needs to determine which differences need to be controlled for in order to generate a valid counterfactual (Howard et al., 2011). Finally, structural models may imbue the researchers priors, such as over what mechanisms are most important, as well as model misspecification in the analysis (Kehoe et al., 2018).

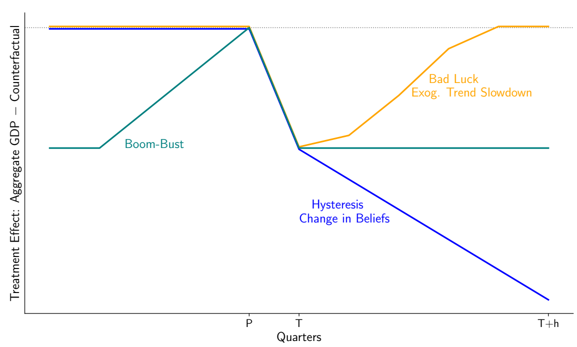
Figure 3.2B shows that regional data can help uncover the underlying counterfactual and thereby discriminate between theories for weak recoveries. For simplicity, suppose that there are two regions, one is completely exposed and one is completely unexposed to the recession. If the two regions are on parallel trends absent the recession, then the unexposed region provides the correct counterfactual for the exposed region.



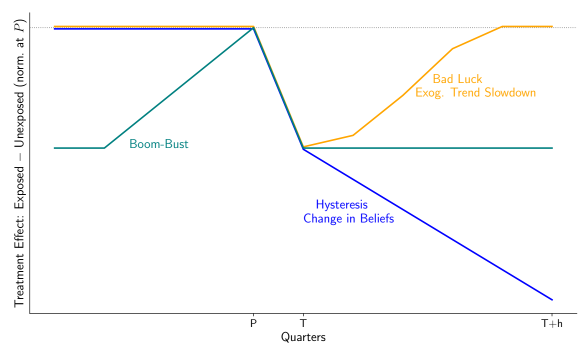
(A) Aggregate Counterfactuals.



(B) Regional Identification.



(C) Aggregate Recession IRF.



(D) Relative Recession IRF.

Figure 3.2. Explanations for Weak Recoveries.

Section 3.2 formalizes identification. In a simple many-good, many-region new Keynesian model, we show that the aggregate effects of the recession can be recovered up to scale from the *relative recession impulse response function*: the relative effect on output in a region that experiences a 1% deeper recession due to its greater exposure to the aggregate business cycle shock compared to another region.

The model shows that the relative recession impulse response function can be estimated with a Bartik IV strategy. In the first stage we predict relative recession depth across regions using a cyclical Bartik that interacts regional industry shares with the cyclical component of national industry growth, and in the second stage we use this instrumented variable to predict relative regional growth before, during, and after the recession. If the cyclical Bartik is uncorrelated with local demand and supply shocks as well as region-specific trends, then this IV strategy recovers the true relative recession impulse response function, and thereby the (scaled) aggregate recession impulse response function.

Figures 3.2C and 3.2D illustrate the identification strategy in the two-region example. The aggregate recession impulse response function in Figure 3.2C is not directly observable and requires modeling a time-series counterfactual. However, the relative recession impulse response function represents the same object and can be feasibly estimated in the data by comparing more and less-exposed regions. In more general settings, general equilibrium forces can scale up or down the regional responses to the aggregate shock, but the aggregate recession impulse response function can still be recovered from the relative recession impulse response function up to scale. As the relative recession impulse response function remains informative about the aggregate recession impulse response, it can help discriminate among theories for weak recoveries.

The cross-region approach holds two important advantages in estimating the recession impulse response function over other methods. First, it is much more robust to trend misspecification than time-series or cross-country panel approaches. Aggregate trends are differenced out, so cross-regional comparisons do not require taking a stance on the stationarity of aggregate GDP, which is notoriously difficult to determine (Stock and Watson, 1999). The construction of the cyclical Bartik still requires a model of industry trends, but we show that idiosyncratic measurement error in industry trend growth washes out at the regional level under mild assumptions on regional specialization. This contrasts with time-series or cross-country panel estimators that are almost surely inconsistent under any trend mismeasurement.

Second, regional data holds the country and the recession fixed, which facilitates identifying the dynamics of the recession. For instance, a more persistent aggregate demand shock will induce a more persistent recession across more and less exposed regions, and thus poses no problem for our identification strategy. In a cross-country or in a time-series setting, two recessions may correspond to shocks with different dynamics. Consider the case in which one recession is a -1% decline in output that lasts one year, and the other recession is a -2% decline in output that shrinks to -1% in the following year and then to zero. Comparing these recessions yields a scaled "treatment effect" of -1% in the first year and -1% in the second year, and thus greater persistence than implied by either recession.

In Section 3.3 we construct the cyclical Bartik instrument based on State industry shares from the quarterly State personal income accounts at the beginning of all post-war recessions and combine them with detrended national industry growth rates from the NBER recession peak to the NBER recession trough. We use leave-one-out growth rate to remove any mechanical correlation of the Bartik with local shocks.

We illustrate how the data inform our estimates in Section 3.4 by exploiting the equivalence between the Bartik approach and weighted GMM using exposure shares (Goldsmith-Pinkham et al., 2020). For pre-1990 recessions the Bartik instrument loads most strongly on the manufacturing sector. We show that States with an above-average manufacturing share experience a deeper recession but also a stronger recovery after the national recession trough. We find a similar pattern in the Covid recession, in which an above-average hospitality and food service share predicts both a deeper recession and stronger recovery. In contrast, a higher construction share during the Great Recession or higher professional services and IT share in 2001 implies a weaker recovery despite a deeper recession. More exposed States also experience relatively strong growth prior to the recession. The different recession-recovery patterns suggest that the relative recession impulse response function is different across recessions with weak and strong national recoveries.

We estimate relative recession impulse response functions in Section 3.5. For recessions with a strong national recovery (pre-1990 and 2020), we estimate a deeper V-shape recession in more exposed U.S. States relative to less exposed States. The more exposed States see a larger fall in income, GDP, and employment and a larger increase in unemployment during the recession, but then grow more quickly in the recovery compared to less exposed States. We cannot statistically reject a complete relative recovery in these recessions, though some point estimates suggest modest hysteresis effects.

For recessions with a weak national recovery, 1990-1, 2001, and the Great Recession, we find that more exposed States experience a stronger boom-bust cycle. In the national recovery phase, the more exposed States do not grow more despite their deeper local recession. However, they do grow more quickly prior to the recession. Thus, the relative recession impulse response function displays an inverse V-shape consistent with boom-bust explanations for weak recoveries (Beaudry et al., 2018; Jordà et al., 2013; Rognlie et al., 2018). Again, the point estimates suggest some modest hysteresis effects in the more exposed States relative to less exposed States, but we cannot statistically reject that a recovery to the pre-boom trend.

Our estimates for strong national recoveries support plucking models of the business cycle (Friedman, 1993; Dupraz et al., 2023), in which recessions are a drop in output below the balanced growth path. For the weak national recoveries of 1990-1, 2001, and the Great Recession, our evidence for is consistent with an "upward" pluck, in which the recession is a drop from an unsustainable boom. To the extent that hysteresis effects are present, these are quantitatively similar across the two types of recessions, which suggests that they are not the main driver of the different recovery patterns.

Following Goldsmith-Pinkham et al. (2020) we decompose our relative recession impulse response functions into the relative recession impulse response functions implied by exposure to a particular sector and the associated sectoral weights. This allows us to test whether changes in the relative recession impulse response function across recessions are due to changes in recession trajectories from particular sectoral exposures or changes in sectoral exposure weights. Across different sectoral exposures, we find evidence of V-shaped relative recessions in more exposed States for the pre-1990 recessions and COVID-19, and boom-bust cycles for the 1990-1, 2001, and Great Recession. This suggests that changing recovery patterns extend beyond the manufacturing sector emphasized by Leamer (2021). Furthermore, changes in sectoral weights play a minor role in explaining weak versus strong recoveries.

We then show that distinguishing between the two types of recessions that we identify matters for policy. We first show that a two-region New Keynesian model with a durable good can match the different recession patterns in the data. Regions that produce relatively more of the durable goods are more exposed to the aggregate business cycle, which allows us to compute relative recession impulse response functions in the same way as in the data. Boom-bust shocks predict a weak recovery in the model because the output in more exposed regions is inflated at the onset of the recession and will diverge from output in less exposed regions. Shocks that induce intertemporal trade-offs, such as monetary policy or a temporary pandemic, instead predict a strong recovery as the recession is simply a temporary deviation from the balanced growth path.

We use the model to calculate the long-run path of the natural rate of interest R^*—the real interest rate that closes the output gap. The model predicts that boom-bust recessions are followed by a prolonged slump in R^*, whereas intertemporal shocks raise R^* in the recovery. This follows from the different dynamics of the capital stock: in boom-bust cycles the capital stock is too high and investment too low, so aggregate demand must be supported by a period of prolonged low real interest rates. However, following intertemporal recessions, the capital stock is too low as investment was delayed, which implies that the recovery is supported by pent-up demand. This shows that appropriately calibrating monetary policy in the recovery depends importantly on a correct diagnosis of the recession.

3.2 Model of Regional Exposure to National Recessions

We develop a multi-region, multi-good new Keynesian model, in which regions are heterogeneously exposed to the aggregate business cycle. The main result is Proposition 3.1, which shows that regional variation in business cycle exposure can be exploited to recover the dynamics of the aggregate business cycle shock, and thereby discriminate among theories for weak aggregate recoveries. The model also illustrates key benefits of using regional data for identifying the dynamics of the business cycle shock: We impose only weak restrictions on the relative mismeasurement of cyclical trends, whereas even iid mismeasurement of aggregate trends will induce bias in other approaches. Furthermore, cross-regional variation by definition loads on the common aggregate recession, whereas approaches that leverage cross-recession variation over time or across countries are likely biased if recessions are not identical up to scale.

3.2.1 Households

A representative household maximizes utility

$$\max_{\{C_{g,t}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \left(\prod_{s=0}^{t} \beta_s \right) \left[\sum_{g=1}^{G} \alpha_{g,t} \frac{C_{g,t}^{1-\sigma_g}}{1-\sigma_g} - \chi \frac{L_t^{1+\eta}}{1+\eta} \right]. \tag{3.1}$$

subject to the budget constraint,

$$\sum_{g=1}^{G} P_{g,t} C_{g,t} + B_t = W_t L_t + (1+i_{t-1}) B_{t-1}, \tag{3.2}$$

where $C_{g,t}$ is consumption of good g at time t, $P_{g,t}$ is its price, B_t are bond holdings, L_t is labor, W_t is the nominal wage, and i_t is the nominal interest rate.

We assume that the household only chooses consumption goods. Labor is not a choice variable but determined by labor demand due to sticky wages. In Section 3.2.8 we discuss an extension of the model with hand-to-mouth consumers.

The first order conditions for the consumption goods are,

$$\alpha_{g,t} C_{g,t}^{-\sigma_g} = \lambda_t \frac{P_{g,t}}{P_t}, \qquad g = 1, \dots, G \tag{3.3}$$

$$\lambda_t = \beta_t (1 + r_t) \mathbb{E}_t \lambda_{t+1}, \tag{3.4}$$

where λ_t is the marginal utility of consumption, P_t is an aggregate price index, and $1 + r_t = (1 + i_t) \frac{P_t}{P_{t+1}}$ is the real interest rate.

Heterogeneity in utility curvature σ_g makes goods differentially sensitive to aggregate shocks. Heterogeneity in curvature σ_g also affects the balanced growth path. In this model we neutralize the latter impact by setting

$$\alpha_{g,t} = \alpha_g Z_{g,t}^{\sigma_g - 1}. \tag{3.5}$$

where $Z_{g,t}$ is good-level productivity. This implies that in our model output of good g will grow linearly in $Z_{g,t}$ irrespective of the curvature σ_g. We make this assumption to create a separation between business cycle sensitivity and long-run growth across sectors. In our quantitative model in Section 3.6, durability determines cyclical sensitivity without affecting long-run growth.

3.2.2 Regional Industry Demand

We assume a standard constant-elasticity of substitution demand system across regions for each good,

$$C_{g,t} = \left(\sum_{i=1}^{I} \alpha_{igt}^{\frac{1}{\eta_g}} C_{i,g,t}^{\frac{\eta_g - 1}{\eta_g}} \right)^{\frac{\eta_g}{\eta_g - 1}}, \tag{3.6}$$

which implies that the demand for good g in region i is given by,

$$C_{i,g,t} = \alpha_{igt} \left(\frac{P_{i,g,t}}{P_{g,t}} \right)^{-\eta_g} C_{g,t}. \tag{3.7}$$

We allow the shares α_{igt} to vary over time across regions,

$$\alpha_{igt} = \alpha_{ig} e^{\hat{\alpha}_{it}}, \tag{3.8}$$

and we discuss the relevance of the region-specific demand shocks $\hat{\alpha}_{it}$ for identification below.

3.2.3 Firms, Technology, Prices, and Market Clearing

Output of good g in region i, $Y_{i,g,t}$, is produced by a representative firm using linear technology,

$$Y_{i,g,t} = Z_{g,t} L_{i,g,t}, \tag{3.9}$$

where $Z_{g,t}$ is good-level productivity and $L_{i,g,t}$ is labor used in region i for good g. Sectoral productivity $Z_{g,t}$ grows according to an ideas production function,

$$Z_{g,t} = \theta_g X_{g,t} Z_{g,t-1}^{1-\xi} Y_{g,t-1}^{\xi}, \tag{3.10}$$

where $X_{g,t}$ is an exogenous productivity trend, the term $Y_{g,t-1}^{\xi}$ captures hysteresis effects of output on productivity when $\xi > 0$, and θ_g is set such that $X_{g,t}$ is the gross productivity growth rate when there are no output fluctuations. We define the exogenous productivity path with

$$\tilde{Z}_{g,t} \equiv X_{g,t} \tilde{Z}_{g,t-1}. \tag{3.11}$$

Firms are perfectly competitive and set prices equal to marginal cost,

$$P_{i,g,t} = \frac{W_{i,g,t}}{Z_{g,t}}. \tag{3.12}$$

We assume wages are indexed to an aggregate productivity index,

$$Z_t = \prod_{g=1}^{G} Z_{g,t}^{\alpha_g}, \tag{3.13}$$

and aggregate prices P_t, but are otherwise insensitive to the business cycle,

$$W_{i,g,t} = \bar{W} P_t Z_t. \tag{3.14}$$

This assumption fulfills two objectives. First, over the long-run wages are consistent with a balanced growth path with full labor mobility across sectors. Second, in the short-run, prices and wages insensitive to the business cycle,

$$\frac{P_{i,g,t}}{P_t} = \frac{P_{g,t}}{P_t} = \bar{W} \frac{Z_t}{Z_{g,t}}. \tag{3.15}$$

In Section 3.2.8 we discuss an extension in which prices respond endogenously to the local business cycle.

Regional goods market clear when output equals consumption,

$$Y_{i,g,t} = C_{i,g,t}. \tag{3.16}$$

3.2.4 Equilibrium

Given exogenous known sequences for the real interest rate $\{r_t\}_{t=0}^{\infty}$, the discount factor $\{\beta_t\}_{t=0}^{\infty}$, expectations of long-run marginal utility of wealth, $\{\mathbb{E}_t \lambda_\infty\}_{t=0}^{\infty}$, regional demand $\{\hat{\alpha}_{it}\}_{t=0}^{\infty}$, and sectoral productivity $\{\{X_{g,t}\}_{g=1}^{G}\}_{t=0}^{\infty}$ the equilibrium is a sequence of $\{\lambda_t, Z_t, Z_{g,t}, C_{g,t}, \frac{P_{g,t}}{P_t}, C_{i,g,t}, Y_{i,g,t}, \frac{P_{i,g,t}}{P_t}\}$ that satisfy (3.3)-(3.16).

We log-linearize this equilibrium around a "balanced growth" equilibrium with the same productivity paths $\{\{Z_{g,t}\}_{g=1}^{G}\}_{t=0}^{\infty}$ and all other sequences at steady state. We denote the outcomes of the balanced growth equilibrium by $\bar{X}_{igt}$ for any variable X, lower case letters are logs $x = \ln X$, and the outcomes in the log-linearized equilibrium by $\hat{x}_{igt} = x_{igt} - \bar{x}_{igt} = \ln X_{igt} - \ln \bar{X}_{igt}$.

3.2.5 Aggregate Business Cycle

We define the aggregate recession impulse response function as the sum of the deviation from the balanced growth path and any hysteresis effects to the balanced growth path,

$$\hat{y}_t + \xi \sum_{s=0}^{t-1} \hat{y}_s = -\left(\sum_{g=1}^{G} \frac{\alpha_g}{\sigma_g}\right)\left[\hat{\lambda}_t + \xi \sum_{s=0}^{t-1} \hat{\lambda}_s\right]. \tag{3.17}$$

The Lagrange multiplier can be solely expressed in terms of the exogenous sequences of the equilibrium,

$$\hat{\lambda}_t = \frac{1}{1-\xi \sum_{g=1}^{G} \frac{\alpha_g}{\sigma_g}} \sum_{s=0}^{\infty} \left(\frac{1}{1-\xi \sum_{g=1}^{G} \frac{\alpha_g}{\sigma_g}}\right)^s (\hat{\beta}_{t+s} + \hat{r}_{t+s}) + \lim_{T\to\infty} \left(\frac{1}{1-\xi \sum_{g=1}^{G} \frac{\alpha_g}{\sigma_g}}\right)^T \mathbb{E}_t \hat{\lambda}_T. \tag{3.18}$$

The growth rate of output is the sum of the growth rate of the business cycle and trend growth,

$$\Delta^h y_{t+h} = \underbrace{\Delta^h \hat{y}_{t+h} + \xi \sum_{s=0}^{h-1} \hat{y}_{t+s}}_{\text{Aggregate Recession Impulse Response Relative to t}} + \underbrace{\Delta^h \tilde{z}_{t+h}}_{\text{Trend Growth}}. \tag{3.19}$$

where Δ^h is an h-period difference.

We next show how we can recover the aggregate recession impulse response function in (3.19) using regional variation.

3.2.6 Regional Bartik Structure

Combining market clearing (3.16), regional demand (3.7), and sticky relative prices (3.15) implies that regional output of good g is proportional to aggregate output of good g and regional demand shifters $\hat{\alpha}_{it}$,

$$Y_{i,g,t} = \alpha_{igt} \left(\frac{P_{i,g,t}}{P_{g,t}} \right)^{-\eta_g} Y_{g,t} = \alpha_{ig} e^{\hat{\alpha}_{it}} Y_{g,t}. \tag{3.20}$$

This setting therefore implies a Bartik structure for local real output growth,

$$\underbrace{\Delta^h y_{i,t+h}}_{\text{Local Output Growth}} = \underbrace{\Delta^h \hat{\alpha}_{i,t+h}}_{\text{Local Demand Shifter}} + \underbrace{\sum_{g=1}^{G} \mu_{i,g,t} \Delta^h y_{g,t+h}}_{\text{Bartik Predicted Growth, } y^B_{i,t,t+h}}, \tag{3.21}$$

where $\mu_{i,g,t} = \frac{P_{i,g,t} Y_{i,g,t}}{P_{i,t} Y_{i,t}}$ is the nominal output share of good g in region i at t.

We can further decompose Bartik predicted growth $y^B_{i,t,t+h}$ into a business cycle $y^{BC}_{i,t,t+h}$ and a trend component $y^{BT}_{i,t,t+h}$,

$$y^B_{i,t,t+h} \equiv y^{BC}_{i,t,t+h} + y^{BT}_{i,t,t+h}, \tag{3.22}$$

where

$$y^{BC}_{i,t,t+h} = \underbrace{\frac{\sum_{g=1}^{G} \mu_{i,g,t} \frac{1}{\sigma_g}}{\sum_{g=1}^{G} \mu_{g,t} \frac{1}{\sigma_g}}}_{\text{Relative Regional Business Cycle Exposure}} \times \underbrace{\left(\Delta^h \hat{y}_{t+h} + \xi \sum_{s=0}^{h-1} \hat{y}_{t+s} \right)}_{\text{Aggregate Recession Impulse Response Relative to t}}, \tag{3.23}$$

$$y^{BT}_{i,t,t+h} = \sum_{g=1}^{G} \underbrace{\mu_{i,g,t}}_{\text{Regional Trend Exposure}} \times \underbrace{\Delta^h \tilde{z}_{g,t+h}}_{\text{Sectoral Trend}}. \tag{3.24}$$

The two Bartik components have different structures. The business cycle Bartik loads on aggregate business cycle shocks based on the covariance of sectoral weights and sectoral sensitivity. The trend Bartik loads on sectoral trends based on sectoral weights alone.

3.2.7 Relative Recession Impulse Response Function

We exploit the Bartik structure to infer the dynamics of the aggregate business cycle. Let the national business cycle peak be $t = P$ and the recession trough be $t = T > P$. We estimate an IV regression, in which we predict local output growth relative to the peak $\Delta^{T-P+h} y_{i,T+h} = y_{i,T+h} - y_{i,P}$ using local recession depth $\Delta^{P-T} y_{i,P} = y_{i,P} - y_{i,T}$, instrumented by the Bartik predicted recession depth, $y^{BC}_{i,P,T}$, while controlling for the Bartik predicted growth relative to the peak, $y^{BT}_{i,P,P+h}$,

$$\text{First Stage:} \qquad \Delta^{P-T} y_{i,P} = \gamma_0 + \gamma_1 y^{BC}_{i,P,T} + \gamma_2 y^{BT}_{i,P,P+h} + \nu_{i,T}. \tag{3.25}$$

$$\text{Second Stage:} \qquad \Delta^{T-P+h} y_{i,T+h} = \delta_h + \beta_h \widehat{\Delta^{P-T} y_{i,P}} + \zeta_h y^{BT}_{i,P,P+h} + \varepsilon_{i,P+h}. \tag{3.26}$$

In what follows we call $\{\beta_h\}_h$ the *relative recession impulse response function*, because it measures the relative behavior of output of a region that experiences a 1% deeper recession compared to another region.

We next state the key aggregation result of this section, that the relative recession impulse response function can recover the aggregate business cycle (3.19) up to scale.

Proposition 3.1. *If Bartik predicted recession depth is uncorrelated with local demand shocks* $Cov_i[y^{BC}_{i,P,T}, \Delta^h \hat{\alpha}_{i,P+h}] = 0$*, then the IV estimator of the relative recession impulse response function recovers the aggregate recession impulse response function scaled by its size during the recession,*

$$\beta^{IV}_h \to - \underbrace{\frac{\Delta^{T+h-P} \hat{y}_{T+h} + \xi \sum_{s=0}^{T+h-P-1} \hat{y}_{P+s}}{\Delta^{T-P} \hat{y}_T + \xi \sum_{s=0}^{T-P-1} \hat{y}_{P+s}}}_{\text{Scaled Aggregate Recession Impulse Response relative to } P}. \tag{3.27}$$

3.2.8 Extensions

In Appendix C.1, we extend the model to include regional hand-to-mouth consumers, prices that respond to the local recession, and endogenous monetary policy. We show that in each case Proposition 3.1 continues to hold unchanged. Intuitively, each of these extensions either scales up or down the effect of the aggregate business cycle shock on regions, which in turn scales the first-stage and reduced-form estimates, or it rescales the aggregate business cycle shock itself. But since our IV estimator normalizes the scale to a 1% relative recession, we recover exactly the same (scaled) aggregate recession impulse response function.

3.2.9 Mapping Relative Recession Impulse Response Functions to Theories

Proposition 3.1 makes clear that our approach is able to discriminate between theories for slow recoveries to the extent that these differ in their scaled aggregate recession impulse response function, which we recover via the relative recession impulse response function β_h. Consider the case of an AR(1) beta-shock and an AR(1) monetary policy shock with identical persistence ρ, and assume there is no hysteresis $\xi = 0$. Then the recession trough immediately follows the peak $T = P + 1$. The relative recession impulse response function recovers the aggregate business cycle persistence:

$$\beta_h = \begin{cases} -\rho^h & \text{if } h \geq 0 \\ 0 & \text{if } h < 0. \end{cases} \tag{3.28}$$

However, the relative recession impulse response function is not informative about the size or identity of the shock. For example, monetary policy shocks and beta shocks with persistence ρ both predict this relative recession impulse response function regardless of their size.

One theory of slow recoveries is that these recessions coincided with a general slowdown in productivity growth (Fernald et al., 2025). Our estimates will be unaffected by simultaneous trend shocks if regional business cycle exposure is uncorrelated with aggregate trends. In this case, if the underlying aggregate business cycle shock is the same across recessions with weak and strong aggregate recoveries, we will recover the same relative recession impulse response function β_h even when the aggregate dynamics are very different.

Another view is that slow recoveries are due to bad luck, in the sense that the economy continues to be subject to negative shocks in the recovery phase (Galí et al., 2012). One interpretation is that the bad luck manifests itself in a more persistent business cycle shock, which implies a slower recovery. In this case, we should observe that the relative recession impulse response function β_h is more persistent in recessions with slow recoveries, but eventually converges to zero as bad luck goes away.

A third view is that boom-bust cycles give rise to slow recoveries. In this framework we can capture a boom through expectations of higher permanent income relative to the balanced growth path, $\mathbb{E}_t\hat{\lambda}_\infty < 0$. For illustration we make the dynamics simple: Households form these optimistic expectations in one period, $\hat{\lambda}_t = \mathbb{E}_t\hat{\lambda}_\infty < 0$, and learn that these expectations are wrong the following period, $\hat{\lambda}_{t+1} = \mathbb{E}_{t+1}\hat{\lambda}_\infty = 0$. These periods then correspond to the recession peak and trough respectively. The relative recession impulse response function is

$$\beta_h = \begin{cases} -1 & \text{if } h \geq 0 \\ 0 & \text{if } h = -1 \\ -1 & \text{if } h < -1. \end{cases} \tag{3.29}$$

That the relative recession impulse response function remains at -1 after the recession implies that regions with a deeper recession do not experience a stronger recovery. Recoveries look weak in this case because the boom represents an unsustainable deviation from the balanced growth path. Importantly, this theory predicts that regions with a deeper recession should grow relatively more prior to the recession.

Now consider the hysteresis case. For simplicity we assume the underlying shock is an AR(1) shock to beta or the real interest rate at $t = 0$. The relative recession impulse response function β_h incorporates the hysteresis effect:

$$\beta_h = \begin{cases} -\rho^h - \xi \frac{1-\rho^{T-P-1+h}}{1-\rho} & \text{if } h \geq 0 \\ 0 & \text{if } h < 0. \end{cases} \tag{3.30}$$

The relative recession impulse response function converges to $-\frac{\xi}{1-\rho}$ because hysteresis causes a permanently lower level of output. Thus, similar to the boom-bust scenario, hysteresis predicts that relative recoveries look weak. In contrast to the boom-bust cycle, there is no pre-trend in the hysteresis relative recession impulse response function.

3.2.10 Identifying Assumptions

The model highlights the necessity of an identification strategy for the relative recession impulse response function, as OLS estimates are almost certainly biased. In particular, local demand shocks $\hat{\alpha}_{it}$ reduce both regional recession depth $\Delta^{P-T} y_{i,P}$ and increase regional growth $\Delta^{T-P+h} y_{i,T+h}$ biasing the relative recession impulse response function β_h downward. The identification problem is reminiscent of cross-country panel approaches, which must assume that country-specific shocks are uncorrelated with recession intensity.

The Bartik approach identifies the relative recession impulse response function if States with greater exposure to the national cycle due to their industry structure do not systematically experience either positive or negative local shocks,

$$\mathbb{C}\text{ov}_i[y^{BC}_{i,P,T}, \Delta^h \hat{\alpha}_{i,P+h}] \propto \mathbb{C}\text{ov}_i[\sum_{g=1}^{G} \mu_{i,g,t} \frac{1}{\sigma_g}, \Delta^h \hat{\alpha}_{i,P+h}] = 0. \tag{3.31}$$

This is the standard Bartik assumption. We remove any mechanical correlation between the Bartik instrument and local shocks by constructing leave-one-out industry growth rates.

We also remove any correlation between the cyclical exposure and exposure to industry trends by controlling for the trend Bartik $y^{BT}_{i,P,P+h}$ in the first and second stage. This is important in practice because sectors such as manufacturing is known to be both cyclically sensitive and in secular decline.

3.2.11 Comparison with Other Approaches

Regional identification of the relative recession impulse response function conveys two important advantages compared to time-series or cross-country panel identification of the aggregate recession impulse response function. First, the construction of the cyclical Bartik does not require taking a stance on the aggregate trend. In practice, it is notoriously difficult to determine whether aggregate GDP is trend-stationary or a random walk (Stock and Watson, 1999), and aggregate trend extrapolation is very sensitive to timing decisions (Ball, 2015).

The construction of the cyclical Bartik does require taking a stance on industry trends, but we show next that it is much more robust to trend misspecification than other approaches. Denote the measurement error in industry-level trend at time t by $\bar{e}_{g,t}$. In Appendix C.2 we show that we continue to identify the relative recession impulse response function if variation in measurement error across goods g is orthogonal to the output growth of the typical good-g producer over the recovery phase, $Cov_g[\Delta^{T-P}\bar{e}_{g,P}, \sum_i \mu_{i,g,t}\Delta^{T+h-P}y_{i,T+h-P}] = 0$.[2] This imposes a restriction on the joint distribution of the covariance structure of measurement error and regional exposures. For example, this restriction is satisfied if variation in measurement error across goods is as good as randomly assigned and regional production is sufficiently diversified in the limit $E_g\left(\sum_i \mu_{i,g,t}^2\right) \to 0$.[3] Consistent with this assumption, Foerster et al. (2022) document that variation in long-run TFP growth across sectors is almost entirely idiosyncratic.[4] Appendix C.2 also provides alternative conditions that allow for correlated measurement error across goods.

[2] Any common trend measurement error across industries is differenced out in the regional regressions (3.25)-(3.26), and thus poses no threat to identification.

[3] These assumptions are equivalent to Assumption 2 in Borusyak et al. (2022). As in their work, our asymptotics implicitly assume that $G(I) \to \infty$ as $I \to \infty$.

[4] Specifically, the factor structure for sector TFP growth explains less than 10% of the variation for each sector, and is statistically insignificant in all cases (Foerster et al., 2022, Appendix Table A1). In contrast, Foerster et al. (2011) document the business cycle fluctuations in industrial production are well explained by a factor model, which is consistent with our model.

Importantly, we do not have to restrict the serial correlation in measurement error in any way. This is because these assumptions imply that good-level trend measurement error washes out asymptotically at the regional level, and thereby justify estimating relative recession impulse response functions at the regional rather than at the industry level.

These assumptions also contrast with the much stronger assumption required in a time-series or cross-country panel approach, where one must impose that measurement error in the estimated aggregate trend be orthogonal to output growth across recession episodes j, $Cov_j[\Delta^{T_j-P_j}\bar{e}_{P_j}, \Delta^{T_j+h-P_j}y_{T_j+h}] = 0$. As we show in Appendix C.2, any stochastic component in aggregate trend growth, such as a productivity slowdown that coincides with recession episodes, will result in a non-zero covariance between the measurement error term and post-recession output growth. Thus, recovering the aggregate recession impulse response function from time-series variation requires trend growth to be either deterministic or measured without error.

Second, regional data holds the country and the recession fixed, which facilitates identifying the dynamics of the recession. For instance, a more persistent aggregate demand shock will induce a more persistent recession across more and less exposed regions, so that our identification strategy will uncover the more persistent dynamics in the cross-section. In a cross-country or in a time-series setting, two recessions may correspond to shocks with different dynamics. Consider the case in which one recession is a -1% decline in output that lasts one year, and the other recession is a -2% decline in output that shrinks to -1% in the following year and then to zero. Comparing these recessions yields a scaled "treatment effect" of -1% in the first year and -1% in the second year, and thus greater persistence than implied by either recession.

The main disadvantage of the regional approach is that is estimates the relative recession impulse response function rather than the aggregate recession impulse response function. Within the simple model in this section, these correspond exactly to each other. In our quantitative model in Section 3.6, this correspondence does not hold exactly, but we show that the relative recession impulse response function remains informative about the aggregate recession impulse response function. In the spirit of Nakamura and Steinsson (2018), it is an identified moment that helps discipline aggregate theories of slow recoveries.

3.3 Data

Our identification strategy requires business-cycle frequency data that is disaggregated by both location and industry for the entire post-war period. We use the quarterly BEA State personal income data, which is disaggregated by industry since 1948. The industry detail increases from 12 SIC sectors from 1948-57 to 14 SIC sectors from 1957-2001 and to 24 NAICS sectors from 1997 onward. Our outcomes are quarterly State personal income, quarterly State unemployment rates from Fieldhouse et al. (2024), quarterly State employment data from the BLS,[5] and, post-2005, quarterly State-level GDP from the BEA. We deflate all nominal variables using the national GDP deflator, and convert all data (except unemployment) to per capita using State population estimates from the U.S. Census linearly interpolated to a quarterly frequency.

When using personal income data we aim to closely proxy for local private output. We therefore exclude the government sector from personal income. We also exclude the finance, insurance, and real estate (FIRE) sector and agriculture and mining. For these sectors we find the largest divergence between local GDP growth and local personal income growth, likely because the geography of ownership and production are more often different in these sectors than in other sectors.

We follow the standard NBER dating for recession peaks and troughs, except that we combine the 1980 and 1981-2 recession into a single episode. We detrend both State recession depth and national industry growth using NBER-peak to NBER-peak growth rates. This gives us (1) a slightly stronger first stage and (2) puts more weight on sectors that experience a sharp change in growth rates in a recession relative to sectors that grow fast or slow due to an underlying trend. For the 1948-9 recession and for the SIC-NAICS switch in 1997 we do not have industry

[5]We seasonally adjust the raw monthly employment data using the Census X-13 algorithm up to February 2020. We then apply the last year of seasonal factors to the remaining sample. This procedure avoids contamination of the seasonal factors by the sharp Covid recession. We then aggregate the seasonally adjusted data to a quarterly frequency.

data from the prior peak, in which case we proxy the trend using the average growth rate from the available data up to the current peak.[6]

3.4 Illustration of the Empirical Approach

Goldsmith-Pinkham et al. (2020) show that estimation using the Bartik instrument (3.23) is equivalent to weighted GMM using the shares s_{i,k,P_j} as instruments.[7] We therefore illustrate the empirical approach based on the sectoral shares that receive the highest weight in the Bartik instrument. Specifically, in Section 3.5.4 we show that the Bartik instrument loads most strongly on the manufacturing share in pre-1990 recessions, on the construction share in the Great Recession, on the accommodation and food sector during COVID-19, and on IT and professional services in the 2001 recession.

Figure 3.3 plots the time path of average per-capita State personal income relative to the national recession peak, when sorted by the key sources of variation in the Bartik instrument. The top left panel sorts U.S. States by their manufacturing share in each recession, and then pools across all pre-1990 recessions.[8] The average recession begins five quarters before the trough, when the high and low manufacturing groups cross at 0 on the y-axis. The difference between the two groups at the NBER trough (0 on the x-axis) therefore measures the relative recession depth. The high manufacturing States decline by 3.4% more during the recessions relative to the low manufacturing States. In other words, the high manufacturing States are more exposed to the national recession.

After the recession trough, high manufacturing areas grow more quickly than low manufacturing areas, which indicates that they catch up from a deeper recession. The high manufacturing areas grow an additional 1.8% in the ten quarters of the recovery, making up more

[6]The Personal Income files switch to a more detailed SIC classification in 1957. We aggregate these to the coarser historical classification for the purpose of constructing peak-to-peak growth rates for the 1957-8 recession.

[7]Adao et al. (2019b) and Borusyak et al. (2022) also show that identification can also be based on exogeneity of the sectoral shocks rather than the sectoral shares. However, their approach requires a large number of independent sectoral shocks, which is not applicable to our setting.

[8]We group States on whether they are above or below the mean share separately for each recession.

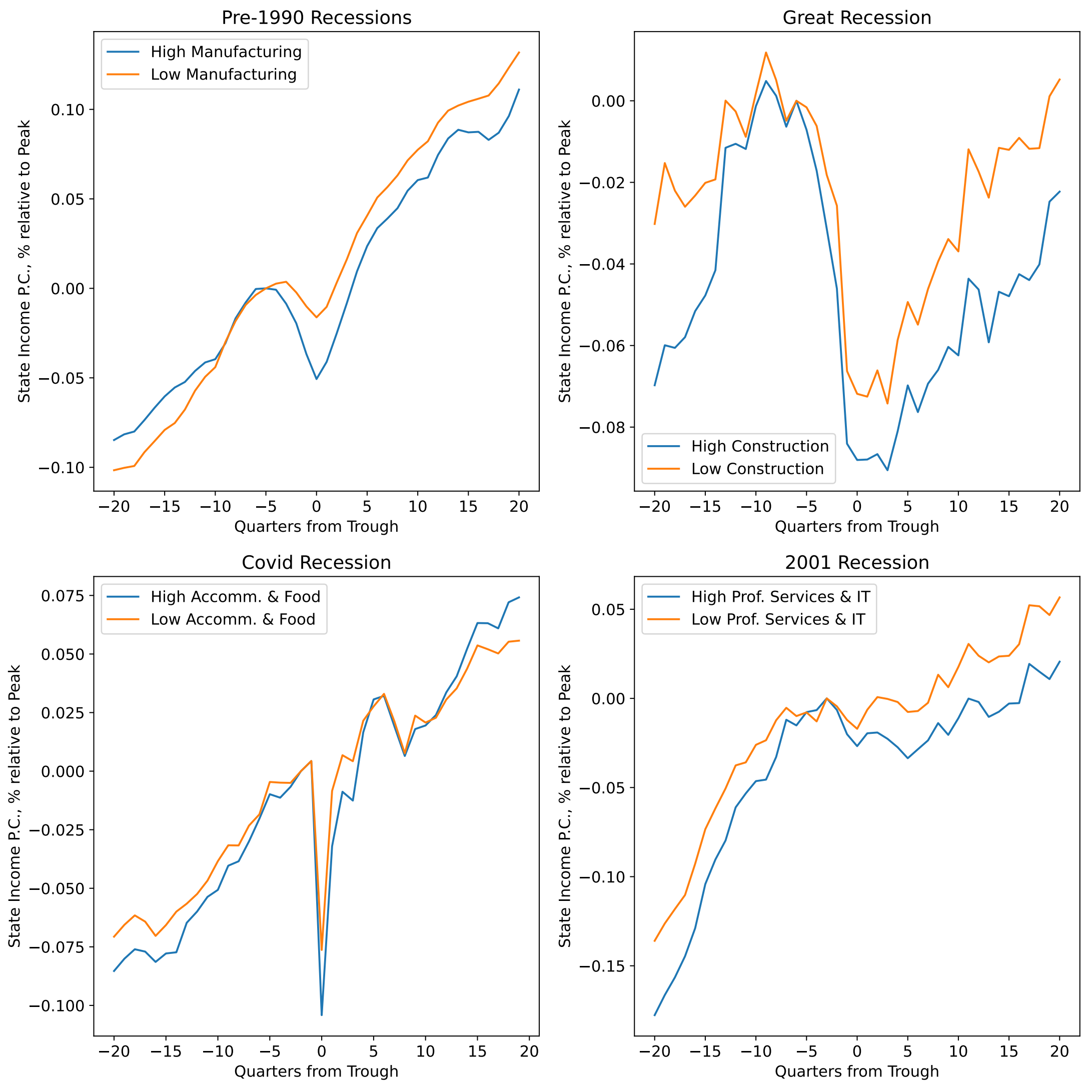


Figure 3.3. Outcomes by Key Sources of Variation of the Bartik Instrument.

than half of the relative recession depth. The top left panel also shows that high-manufacturing States grow 0.11% less per quarter relative to low-manufacturing States before the recession. Accounting for this relative trend implies that the relative recovery is complete after 10 quarters.[9]

The pre-1990 recession plot also illustrates that our cross-sectional recession elasticities are robust to misspecification of the aggregate trend, unlike a time-series analysis. Both the low and high manufacturing areas are 2.4% below their pre-recession trend 10 quarters after the recession trough. But this may reflect that the aggregate pre-recession trend was too high for both groups. Blanchard et al. (2015) show that time-series estimates of output gaps are very sensitive to assumptions about the aggregate pre-recession trend. Our analysis does not require taking a stance on the aggregate pre-recession trend.

In the top right panel of Figure 3.3 we show that U.S. States loading more on the construction sector in the Great Recession experience both a deeper recession and a weaker recovery. Specifically, high construction exposure predicts a 1.6% deeper recession and a 0.9% weaker recovery in the first ten quarters after the NBER trough. These patterns are very different from the pre-1990 recession, in which more exposed States were catching up after the recession. Another difference is the pre-trend: The more exposed States in the Great Recession grew faster than the less exposed States. These patterns inform our cross-sectional regression that the Great Recession was a boom-bust cycle.

The COVID-19 recession shows a similar pattern to the pre-1990 recessions: States with a steeper recession due to greater exposure to accommodation and food also experience stronger recoveries. The catch-up in levels is nearly complete: the recession is 2.8% deeper in the States with a high accommodation and food share, but the recovery is also 2.6% stronger.

The final subplot on the bottom right shows that for the 2001 recession the patterns are qualitatively similar to the Great Recession: a deeper recession due to a greater loading on IT and professional services implies a weaker recovery. More exposed States also grow more quickly

[9]The relative recovery adjusting for the trend is 1.8% + 15*0.11% = 3.4%, which is equal to the 3.4% relative recession depth.

before the recession, again indicating boom-bust dynamics across States.

In summary, the data suggests that the pre-1990 and the Covid recessions have V-shaped relative recession impulse response functions, in which a deeper recession predicts a stronger recovery. In contrast, exposure to the Great Recession and the 2001 recession predicts an inverse-V shaped relative recession impulse response function: faster growth prior to the recession, followed by a steeper decline and weaker recovery. We next make this finding explicit using regression analysis.

3.5 Results

3.5.1 First Stage

We pool our estimates across the two recession groups: those with weak aggregate recoveries (1990-1, 2001, and 2007-9) and those with strong aggregate recoveries (1948-9, 1953-4, 1957-8, 1960-1, 1969-70, 1973-5, 1980-2, 2020). Pooling helps to statistically distinguish these recession groups and is supported by the similar qualitative recession pattern we document above.

We start with the pooled first stage,

$$\begin{aligned}\Delta^{P_j-T_j} y_{i,P_j} = {} & \gamma_1 y^{BC}_{i,P_j,T_j} I(P_j \in \{1990, 2001, 2007\}) + \gamma_2 y^{BC}_{i,P_j,T_j} I(P_j \in \{\text{Other Recs.}\}) \\ & + \delta_1 y^{BT}_{i,P_j,T_j} I(P_j \in \{1990, 2001, 2007\}) + \delta_2 y^{BT}_{i,P_j,T_j} I(P_j \in \{\text{Other Recs.}\}) \\ & + \boldsymbol{\zeta} \mathbf{X}_{i,j} + \theta_j + \varepsilon_{i,j,h},\end{aligned} \tag{3.32}$$

in which we regress recession depth relative to trend $\Delta^{P_j-T_j} y_{i,P_j} = y_{i,P_j} - y_{i,T_j}$ on Bartik-predicted recession depth y^{BC}_{i,P_j,T_j}. We allow the relationship to vary across recession groups. γ_1 is the elasticity of actual recession depth w.r.t. predicted recession depth in the 1990-1, 2001, and 2007-9 recessions. γ_2 captures the average elasticity for the other recessions. We include recession fixed effects θ_j that isolate within-recession variation, consistent with Proposition 3.1.

Table 3.1. First Stage: Predicting Local Recession Depth.

	Incl. Earn.	All Earn.	Unempl.	Employ.	GDP
Coeff. on Predicted Recession Depth:					
90-1, 01, 07-9	2.795***	2.242***	58.459**	1.993***	4.068***
	(0.750)	(0.656)	(27.188)	(0.487)	(0.956)
Pre-1990 & 2020	1.487***	0.866***	64.896***	0.985***	1.078***
	(0.334)	(0.222)	(9.541)	(0.215)	(0.217)
Bartik Trend Control	Yes	Yes	Yes	Yes	Yes
Local Pre-Trend Control	Yes	Yes	Yes	Yes	Yes
Recession FE	Yes	Yes	Yes	Yes	Yes
F-stat (90-1, 01, 07-9)	13.9	11.7	4.6	16.7	18.1
F-stat (pre-1990/2020)	19.8	15.2	46.3	21.0	24.6
p-val equality	0.114	0.050	0.824	0.061	0.003
State Clusters	100	100	100	100	100
R^2	0.136	0.056	0.137	0.150	0.408
Observations	544	543	490	535	100

Note: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered by state.

We control for Bartik-predicted trend, y^{BT}_{i,P_j,T_j}, separately for each recession group. The control vector $\mathbf{X}_{i,j}$ contains the State's average growth rate from 15 years before the recession to 5 years before the recession, again interacted with recession group. This control addresses any State-specific trends uncorrelated with industry trends. Figure 3.3 shows that differentially exposed States are on different pre-trends, which indicates either differences in underlying trend growth or a boom-bust cycle. We start the trend control five years prior to the recession so as to minimize capturing boom-bust dynamics, and we show that our results are robust to extending this window. Within each Recession Group we cluster by State.[10]

The pooled first stage estimates are shown in Table 3.1. In columns (1) and (2) we show that predicted recession depth based on State income shares and national trends strongly predicts

[10]Double-clustering by both State and recession date and adjusting for small numbers of clusters following Imbens and Kolesar (2016) generally produces smaller standard errors than clustering by State alone. See Figure C.1. This suggests that our Bartik approach largely captures cross-state correlations, consistent with Adao et al. (2019b).

State recession depth in income. For the 1990-1, 2001, and 2007-9 recessions the first stage coefficient for income included in the instrument is $\gamma_1 = 3$ (t-statistic 4.5). This means that a 1% deeper predicted recession translates into a 3% deeper local recession. For the other recessions the first stage coefficient is $\gamma_2 = 1.4$ (t-statistic 4.4). In columns (3) through (5) we measure State recession depth relative to trend with the unemployment rate, employment, and GDP. In each case, our Bartik instrument is a powerful predictor of the State recession depth.

3.5.2 Relative Recession Impulse Response Functions

We next estimate relative recession impulse response functions across recession groups. Within each group, these capture the differential movement of income across more and less exposed States both before and after the national recession.

Specifically, we pool equation (3.26) across recession groups,

$$\begin{aligned} y_{i,T_j+h} - y_{i,P_j} = {} & \beta_{1h} \widehat{\Delta^{P_j - T_j} y_{i,P_j}} I(P_j \in \{1990, 2001, 2007\}) \\ & + \beta_{2h} \widehat{\Delta^{P_j - T_j} y_{i,P_j}} I(P_j \in \{\text{Other Recs.}\}) \\ & + \delta_{h1} y^{BT}_{i,P_j,T_j} I(P_j \in \{1990, 2001, 2007\}) + \delta_{h1} y^{BT}_{i,P_j,T_j} I(P_j \in \{\text{Other Recs.}\}) \\ & + \boldsymbol{\xi} \mathbf{X}_{i,j} + \alpha_{j,h} + \varepsilon_{i,j,h}, \end{aligned} \tag{3.33}$$

where $y_{i,T_j+h} - y_{i,P_j}$ is h-quarter growth in the outcome variable since the recession peak P_j. $\widehat{\Delta^{P_j - T_j} y_{i,P_j}}$ is the instrumented recession depth from the first stage (3.32).

Figure 3.4 plots our estimates for the relative recession impulse response functions, β_{1h} and β_{2h} for $h = -40$ quarters before the recession trough to $h = 20$ quarters after the recession trough. For $h = 0$, the estimate is close to -1 for both recession groups. If there were no trends in our measure of local recession depth, these estimates would be equal to -1 exactly. They are close to -1 because the relative trends are not very important over the short recession duration.

Other than the immediate period prior to the recession trough $h = 0$, the two relative recession impulse response functions look different. For the 1990-1, 2001, and 2007-9 recessions,

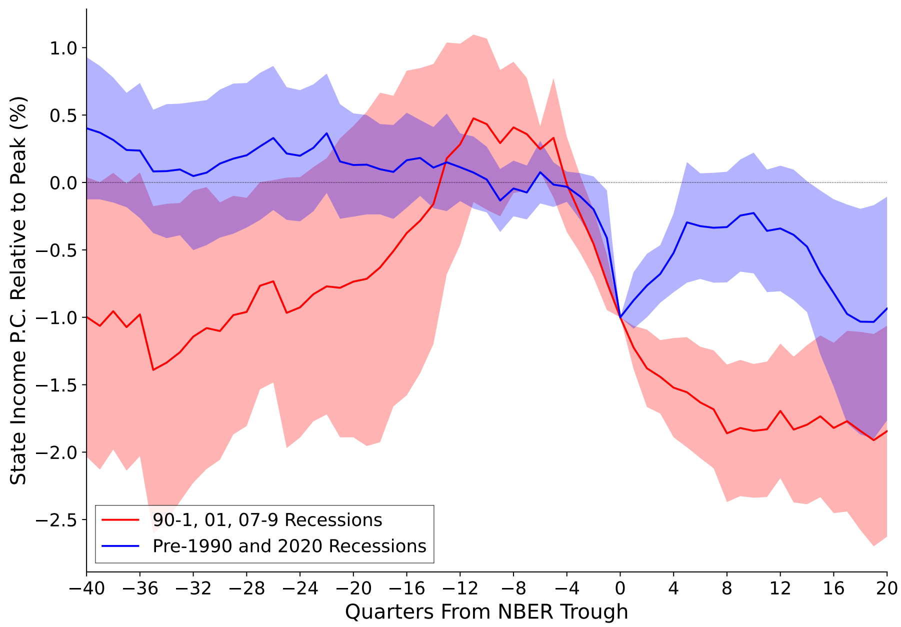


Figure 3.4. IRF of State Personal Income P.C. by Recession Group.

a deeper relative recession predicts a weaker relative recovery. This is reflected in the decreasing estimates of the relative recession effect for $h > 0$. Thus, States that are more exposed to the recession experience relatively weaker growth in the recovery compared to less exposed States.

From $h = -20$ to $h = -10$, the more exposed States grow relatively more quickly compared to the less exposed States in the 1990-1, 2001, and 2007-9 recessions. This is captured by the increase in β_{1h} over this window from roughly -1 to slightly above 0. The strong relative growth in more exposed States is a break from their prior behavior. Going from $h = -20$ back as far as $h = -40$, the more exposed States are growing at a similar pace than the less exposed States. The point estimates for the recovery period are somewhat below the pre-recession trend, which is indicative of modest hysteresis. However, the confidence intervals are also consistent with no hysteresis effects in the 1990-1, 2001, and 2007-9 recessions.

The relative recession impulse response function for the pre-1990 and 2020 recessions is very different. Prior to the recession, the more exposed States are on a slight downward trend relative to the less exposed States, but the point estiamtes are not statistically different from zero. During the recession, the more exposed States experience a sharp relative decline, followed by an almost equally sharp relative recovery. Four quarters after the recesssion trough, we can no longer reject the null that the more exposed States have fully caught up to the less exposed States. The point estimates for the recovery are again consistent with modest hysteresis. [11]

Through the lens of our model in Section 3.6, the relative recession impulse response function for the 1990-1, 2001, and 2007-9 recessions is most consistent with a boom-bust cycle (Beaudry et al., 2018; Jordà et al., 2013; Rognlie et al., 2018), in which more exposed States experience a stronger boom followed by a stronger bust and a return to the pre-boom trend. The relative recession impulse response function for the pre-1990 and 2020 recessions is consistent with plucking models of the business cycle (Friedman, 1993; Dupraz et al., 2023). The more exposed States are plucked down by a shock during the recession, but then return to their prior

[11]The second decline at $h > 12$ quarters reflects a subsequent recession: the peak of the 1960-1 recession follows only eleven quarters after the 1957-8 recession.

trend in the recovery. While we find some evidence for hysteresis effects, their magnitude is very similar across the two recession groups, which suggests that the difference in recovery strength is not driven by a hysteresis mechanism. Instead our evidence suggests that the weak recoveries in the 1990-1, 2001, and 2007-9 recessions relative to the other recessions are due to different types of shocks.

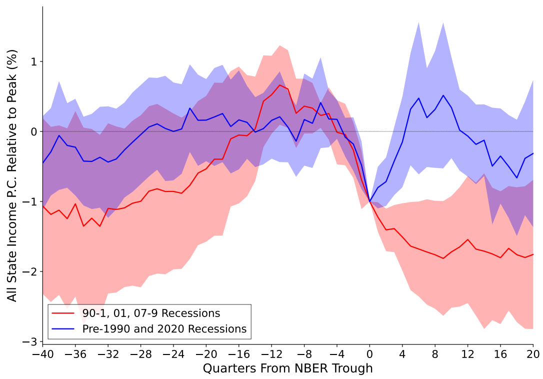
3.5.3 Other Outcomes

Our baseline outcome variable is State income, specifically the components of State income that were included in our Bartik instrument. We next show impulse response functions for four alternative outcomes: all State income (including all sectors, transfers, dividends, and rent), the State unemployment rate, State employment, and for the last two recessions, State GDP. The four impulse response functions are plotted in Figure 3.5.

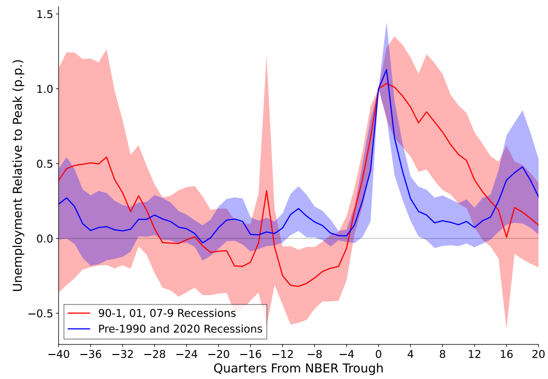
The boom-bust pattern for all State income in the 1990-1, 2001, and 2007-9 recessions is essentially the same as in our baseline estimates, though less precisely estimated (Figure 3.5A). For the pre-1990 and 2020 recessions, the recovery is more complete based on all income and suggests a complete absence of hysteresis effects.

The unemployment impulse response function in Figure 3.5B is noticeably different from the other outcomes. It displays convergence in unemployment rates across States for all recessions, consistent with prior work by Blanchard and Katz (1992) and Yagan (2019). But the speed is very different: For the 1990-1, 2001, and 2007-9 recessions convergence takes four years, whereas for the pre-1990 and 2020 recessions convergence is almost complete after one year.

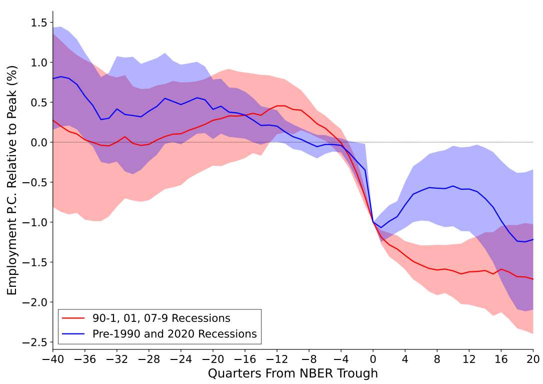
There are no noticeable pre-trends in unemployment for the pre-1990 and 2020 recessions. In contrast, unemployment in more exposed States was relatively low at the peak of the 1990-1, 2001, and 2007-9 recessions, consistent with a relative boom in the labor market. However, this boom is relatively small compared to other outcomes. The relative boom in State income in Figure 3.4 in the five years prior to the recession is almost as large as the relative bust. The relative decline in unemployment in the prior to the recession is only one-quarter as large as the subsequent increase in unemployment. This suggests that boom-bust cycles may have asymmetric effects on labor markets and that unemployment rates are less informative about the boom-bust nature of a recession than other outcomes.



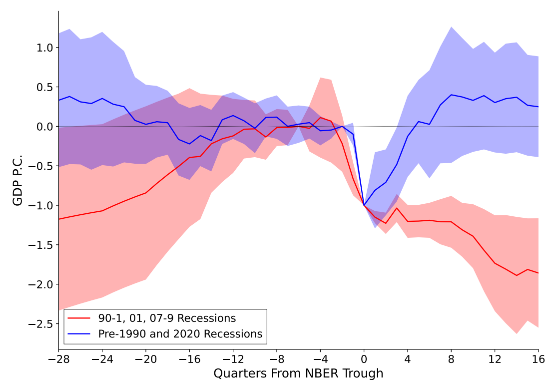
(A) All State Income P.C..



(B) Unemployment Rate.



(C) Employment P.C..



(D) GDP P.C..

Figure 3.5. Recession Impulse Response Function for Other Outcomes.

Figure 3.5C plots the relative recession impulse response function for employment to population ratio. The patterns are qualitatively similar to the income results: a sharp recovery in the pre-1990 and 2020 recessions, and a boom-bust cycle in the 1990-1, 2001, and 2007-9 recessions. But quantitatively there are two notable differences. First, there is weaker evidence that the more and less exposed States are on parallel trends. In particular, for the pre-1990 and 2020 recessions, the more exposed States are on a relative downward trend prior to the recession. For the 1990-1, 2001, and 2007-9 recessions, the pre-trend appears to be first negative before turning positive in the five years prior to the recession. Second, there is stronger evidence for hysteresis effects in employment than in income. The failure of parallel trends does suggest this interpretation should be taken with caution. However, regardless of the interpretation, the hysteresis effects are not different across recession groups, and thus do not explain the differential recovery speed.

The GDP data is more limited since it is only available for the last two business cycles. Yet the patterns are similar to the broader sample for State income. The Covid recession displays a fast recovery to the pre-recession trend. The Great Recession shows a boom-bust cycle. These outcomes corroborate our findings using State income.

For comparison, we plot the relative growth rates of population for the two recession group in Appendix Figure C.3. Population grows relatively more in the more exposed States in the 1990-1, 2001, and 2007-9 recessions, but relatively less in the other recessions. Population growth remains smooth throughout the recession, unlike the economic outcomes in Figures 3.4 and 3.5. Our results extend the findings of Autor et al. (2021) for the Great Recession and support their conclusion that the "estimates for population belie the expectation of a lasting downturn in the CZs with greater exposure to the Great Recession."

3.5.4 Decomposition of the Impulse Response Function

We decompose our Bartik estimates to (1) show what variation is most important and (2) determine whether the changes across recessions come from different behavior within or across industries (or both). Goldsmith-Pinkham et al. (2020) show that the Bartik estimator is equivalent to a weighted GMM estimator, in which the local industry shares s_{i,k,P_j} in (3.23) are the instruments and industry growth rates $\hat{g}_{-i,k,P_j,T_j}$ are the weights. When the industry growth rates are common across locations, then we can decompose the Bartik estimate for β_h in (3.26) as

$$\beta_h = \sum_{k,j} \alpha_{k,P_j} \beta_{h,k,P_j}, \tag{3.34}$$

where β_{h,k,P_j} is the estimate for β_h from instrumenting (3.26) using the industry share s_{i,k,P_j} and α_{k,P_j} is its associated "Rotemberg" weight in the Bartik instrument. In our setting, this decomposition holds only asymptotically since we use leave-one-out growth rates to construct the Bartik instrument. In practice, the estimates from the leave-one-out Bartik and the common growth rate Bartik are very close, so the decomposition remains informative about what sources of variation matter for our estimates.

We focus on the estimates for State income from Figure 3.4. In Figures 3.6A and 3.6B we show how these impulse response functions are informed by exposure to five broad sectoral aggregates that we can consistently define throughout the sample: Construction, manufacturing, transportation and utilities, wholesale and retail trade, and other services. Each estimated impulse response function uses the corresponding local income share of that sector as instrument for local recession depth. When a sector share is largely uninformative about local recession depth, then its corresponding Rotemberg weight α_{k,P_j} will be small and the standard error for β_{h,k,P_j} will be large. We therefore exclude sectors with a weight less than 0.05 from the figure. Appendix Figure C.2 includes all sectors.

The three major sectors that inform recession depth in the 1990-1, 2001, and 2007-9 recession are construction ($\alpha = 0.3$), transportation and utilities ($\alpha = 0.08$), and other services ($\alpha = 0.54$). Figure 3.6A shows that recession exposure based on each of these major sectors predicts a boom-bust cycle. In almost all cases, the pooled estimate is within the 95% confidence band of each of the sectoral exposures. Thus, the boom-bust nature of these recessions is broad-based rather than confined to a particular industry. It also suggests that the boom-bust cycle is unlikely to be due to a correlation of sectoral exposure and local income shocks, as one then would expect the boom-bust cycle to be confined to sectors that are overrepresented in locations subject to those local shocks.

Leamer (2021) argues that weak recoveries in the manufacturing sector help explain weak aggregate recoveries in the 1990-1, 2001, and 2007-9 recessions based on cross-State data. While we also find that manufacturing exposure predicts weaker recoveries in the 1990-1, 2001, and

2007-9 recessions, exposure based on the other sectors yields a similar pattern. Furthermore, these other sectors matter more both in the cross-section and in the aggregate, based on their Rotemberg weights and aggregate sectoral shares. This suggests that an explanation for weaker recoveries based solely on the changing behavior of manufacturing is likely incomplete.

For the national recessions before 1990 and in 2020, exposure is primarily determined by the local income shares in manufacturing ($\alpha = 0.52$), wholesale and retail trade ($\alpha = 0.08$), and other services ($\alpha = 0.37$). Figure 3.6B shows that each of the three sectoral exposures predicts a rapid recovery from the NBER trough in these recessions. There is no evidence of boom-bust cycles. If anything, the more exposed States based on manufacturing and wholesale and retail trade are on a downward-trend relative to less exposed States. The recovery from the recession is consistent with a return to this downward-trend. Service sector exposure predicts no differential trend prior to the recession and a complete recovery thereafter.

Comparing Figures 3.6A and 3.6B, we find that all major sector exposures predict weak recoveries in the 1990-1, 2001, and 2007-9 recessions and strong recoveries in the other recessions. This implies that the weaker recovery dynamics are coming primarily from within-sector changes rather than changes in sectoral weights.

Figures 3.6C and 3.6D show the impulse response functions when the variation is restricted to particular recessions. For both the 2001 and 2007-9 recession the boom-bust pattern is evident and each contains the pooled estimate within their 95% confidence bands. This is despite these recessions loading on different sectors: Cross-State variation in recession depth during the Great Recession is best explained using exposure to the construction and health sectors. For the 2001 recession, exposure to health and information and professional services matters most. The Figure omits the 1990-1 recession in which local industry shares have little predictive power for local recession depth, which is reflected in both a low Rotemberg weight $\alpha = 0$ and wide confidence band (Appendix Figure C.2C).

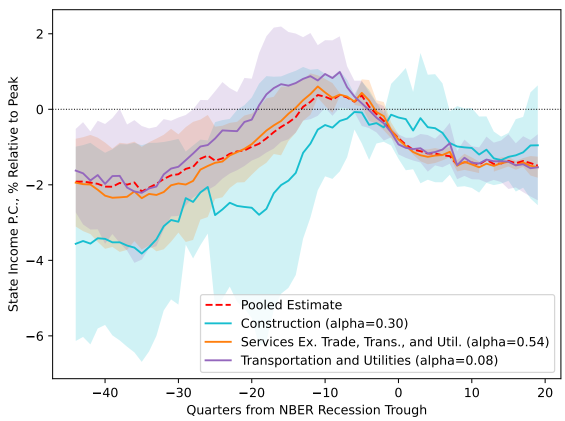
There is no evidence of a boom-bust cycle in the pre-1990 or the Covid recessions in Figure 3.6D. Instead, these recessions show a rapid recovery in the quarters after the NBER trough. States more exposed to the 2020 recession show a complete relative recovery within a year to their pre-recession relative trend. This pattern is primarily informed by exposure to the accommodation and food and the entertainment sectors: States like Nevada and Hawaii contracted sharply during Covid but then rapidly recovered as the economy re-opened. The more exposed States to pre-1990 recessions also grow relatively quickly after the recession trough. In this sense, the pre-1990 recessions look similar to the Covid recession and different from the 1990-1, 2001, and 2007-9 recessions. However, States exposed to the pre-1990 recessions and Covid recession differ in their pre-trends and thus in the extent to which their relative recovery looks complete in levels after the recession. For the pre-1990 recessions, the more exposed States are on a slight downward trend, whereas in the Covid recession the more exposed States are on a slight upward trend prior to the recession.

In short, weaker recoveries following a boom-bust cycle are a broad-based phenomenon across sectors and recessions for the 1990-1, 2001, and 2007-9 recessions. In contrast, the strong recoveries from the pre-1990 and Covid recessions are also broad-based, and they do not display any boom-bust dynamics.

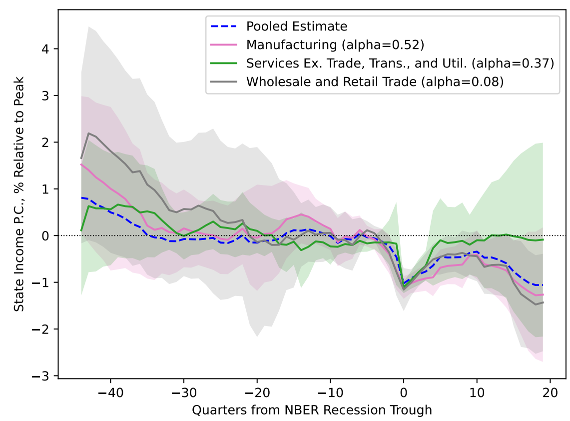
3.5.5 Robustness

We conclude our empirical analysis by showing that our main results are robust to a variety of changes in the construction of the instrument, the set of fixed effects, and the sample. We present these robustness checks in Figures 3.7A to 3.7F. Each figure contains six panels. The left column shows the IRFs for the 1990-1, 2001, and 2007-9 recessions and the right column for the other recessions. The top panel shows how the IRFs change when we vary the instrument (Figures 3.7A and 3.7B), the middle panel varies the fixed effects (Figures 3.7C and 3.7D), and the bottom panel changes the sample (Figures 3.7E and 3.7F). In each case, we re-estimate the first stage and second stage for each variation.

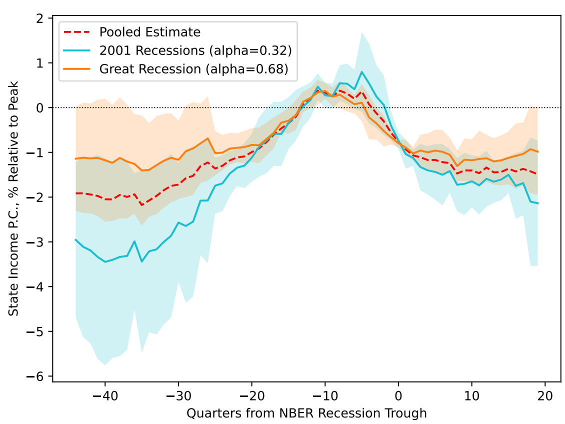
Figures 3.7A and 3.7B show how the IRFs for the 1990-1, 2001, 2007-9 recessions and the other recessions change when we construct the instrument differently. In (2) we construct



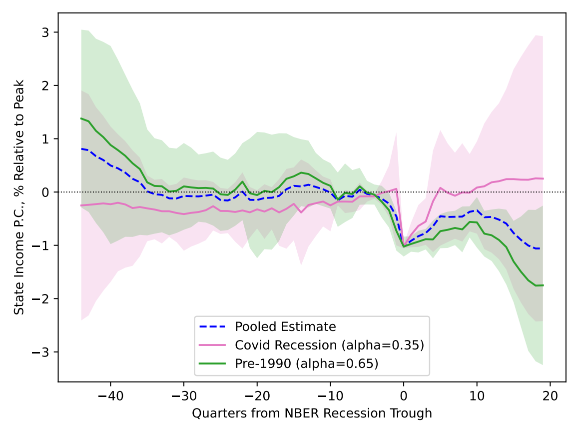
(A) By Industry: 1990-1, 2001, 2007-9.



(B) By Industry: Other Recessions.



(C) By Recessions: 1990-1, 2001, 2007-9.



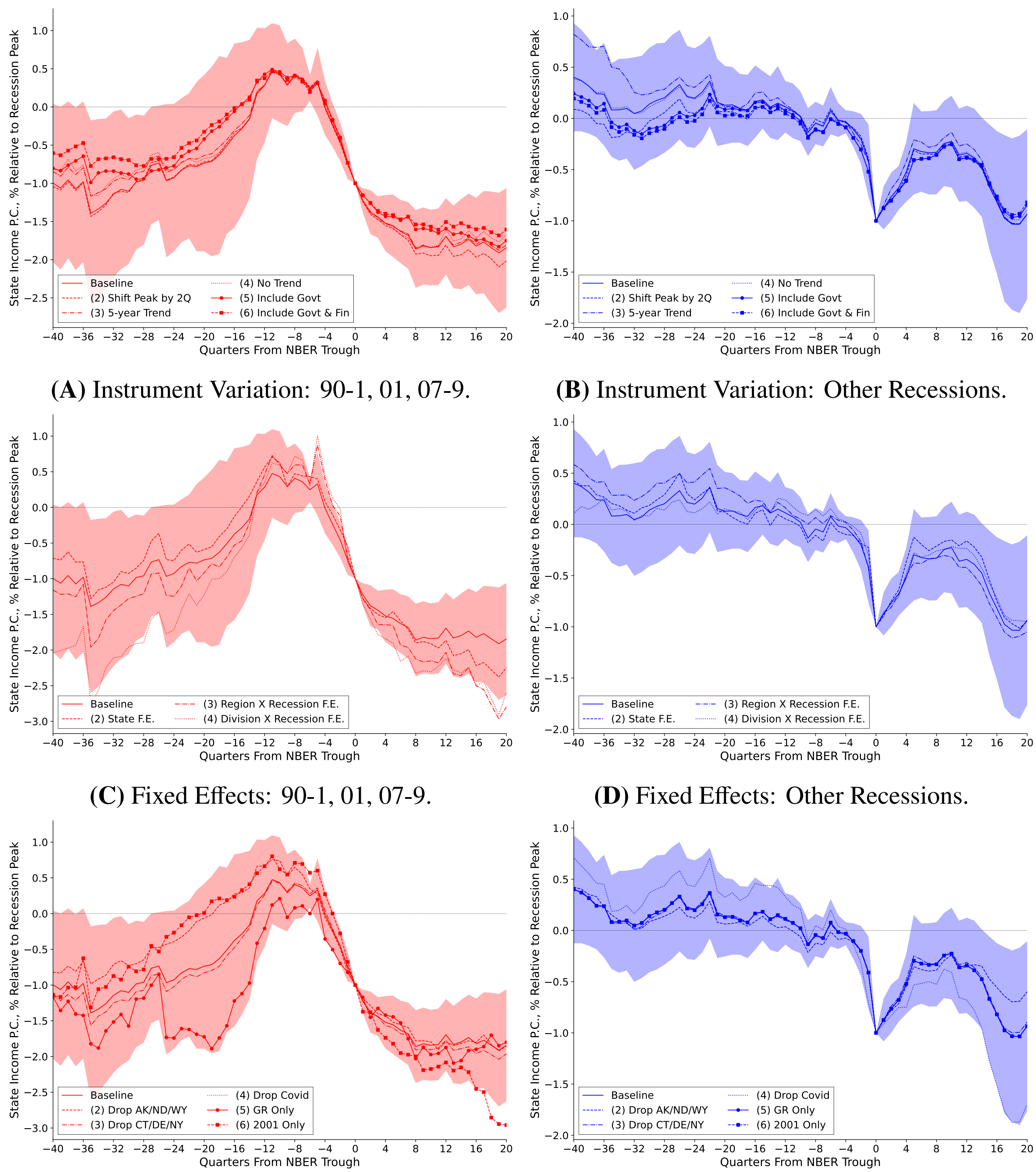
(D) By Recessions: Other Recessions.

Figure 3.6. IRFs of State Income P.C. by Industry and Recession Variation.

the peak-to-peak national industry growth rates two quarters prior to the NBER peak and in (3) we construct it using five-year pretrend from the NBER peak. In (4) we omit detrending local recession depth. Neither of these changes materially affects the impulse response functions. In (5) and (6) we include the government and government and finance sectors in local income shares. This means that as well as comparing across private industries we also compare more cyclical private industries with the less cyclical government sector. The main change is that the boom-bust pattern for the 1990-1, 2001, 2007-9 recessions starts earlier and is a bit more pronounced.

In Figures 3.7C and 3.7D we vary the set of fixed effects to proxy for local shocks correlated with national recession exposure. In (2) we include State fixed effects which absorb any permanent difference across States such as permanently different trends correlated with recession exposure. We find that State fixed effects primarily affect the IRF for the fast recoveries by flattening the pretrend and making the recovery look more complete. In (3) and (4) we use Census region and Census division fixed effects interacted with recession fixed effects. These fixed effects capture local shocks that are regional in nature even if they vary across recessions. They primarily change the IRF for the 1990-1, 2001, 2007-9 recessions by increasing the size of both the boom and the bust.

In Figures 3.7E and 3.7F we show the sensitivity to the sample. In (2) we drop the three most mining-intensive States and in (3) the three most finance-intensive States, the two categories in which we are least confident that the location of income corresponds to production. The differences from our baseline are minimal. (4) drops the Covid recession, which makes the downward-trend in the other recessions more pronounced. (5) and (6) show only the Great Recession and the 2001 recession, respectively. Both display the boom-bust pattern of our baseline.



(A) Instrument Variation: 90-1, 01, 07-9.

(B) Instrument Variation: Other Recessions.

(C) Fixed Effects: 90-1, 01, 07-9.

(D) Fixed Effects: Other Recessions.

(E) Sample Variation: 90-1, 01, 07-9.

(F) Sample Variation: Other Recessions.

Figure 3.7. IRFs Robustness.

3.6 Quantitative Model

The objective of this section is to build a quantitative model that is consistent with the two different types of recessions we document above and then to use that model for policy analysis. The model has two region and two goods, with goods subject to standard New-Keynesian price stickiness. The two-good structure allows for the regions to be differentially exposed to a national recession so we can run the same regressions in the model as we do in the data.

3.6.1 Households

Households maximize expected utility subject to their budget constraint. The expected utility function of household i is given by:

$$\mathrm{E}_{i,0}\sum_{t=0}^{\infty}\beta^t\left[\frac{(C_t)^{1-\sigma}}{1-\sigma}+\psi_t\varepsilon_{\psi t}\frac{{D_{i,t}}^{1-\sigma_d}}{1-\sigma_d}-\varphi_t\frac{{L_t}^{1+\nu}}{1+\nu}\right], \tag{3.35}$$

where C_t is consumption of a composite nondurable good, $D_{i,t}$ is the durable stock of household i and L_t is total labor supply. ψ_t and φ_t are time-varying preference shifters for durable goods consumption and labor supply that keep hours worked and the ratio of nondurable to durable consumption constant along the balanced growth path.[12] $\varepsilon_{\psi t}$ is an exogenous stationary shock to durable preferences that we use to model a boom-bust cycle.

The durable stock D_{it} is specific to the household. We assume that each period only a fraction $(1-\theta^d)$ of households is allowed to re-optimize its durable stock. This formulation, which follows Evans and Ramey (1992) and Orchard et al. (2023), reduces the counterfactually

[12] The process for φ_t is such that it implies constant hours along a balanced growth path:

$$\varphi_t = (C_t)^{-\sigma}Z_t \tag{3.36}$$

where Z_t is the level of productivity of the aggregate economy. The taste for durables, ψ_t, contains a productivity component as well as an exogenous stationary process $\varepsilon_{\psi t}$:

$$\psi_t = \bar{\psi}Z_t^{\sigma-\sigma^d}. \tag{3.37}$$

high sensitivity of durable consumption to interest rate changes present in standard models, while preserving tractability. Nondurable consumption C_t is identical between households due to consumption insurance and labor supply L_t is identical because it is set by labor unions.

Expectations are specific to household i due to information stickiness as in Auclert et al. (2020) and Carroll et al. (2020). The current state at t is common knowledge, so households never violate their budget constraint. There is uncertainty about the future path of aggregate shocks and households update their expectations with i.i.d. probability $1-\theta$ in each period.

The aggregate budget constraint and the law of motion for the aggregate durable stock are given by:

$$A_t = (1+i_{t-1})A_{t-1} + W_t L_t - P_t C_t - \iota P_t D_t - P_t^x X_t \tag{3.38}$$

$$D_t = (1-\delta)D_{t-1} + X_t \tag{3.39}$$

$$D_t = \int_0^1 D_{i,t} di, \tag{3.40}$$

where P_t is the price index for nondurable goods, A_t are nominal bond holdings, i_t is the nominal interest rate in terms of nondurable goods, $W_t L_t$ is nominal labor income, D_t is the aggregate durable stock, P_t^x is the price index of durable goods , X_t is the composite durable good and ιD_t are operating costs from using the durable goods (e.g. gas, insurance), which we assume are denominated in units of nondurable goods.

Durable purchases, X_t, include maintenance costs, given by $\chi\delta D_{t-1}$, and new durable purchases, I_t. Maintenance costs require households to pay a portion χ of depreciation as maintenance expenditures in each period. This yields:

$$X_t = \chi\delta D_{t-1} + I_t \tag{3.41}$$

$$I_t = D_t - (1-(1-\chi)\delta)D_{t-1}. \tag{3.42}$$

Households optimally choose the schedule for $\{C_{it}, D_{i,t}, A_{it}, X_{it}\}_{t=0}^{\infty}$ that maximizes their utility given their information set. Labor supply is not chosen directly by the household, but by regional labor unions which we describe later on. The separability of preferences allows us to break up the household problem in two steps. We first solve for the optimal choices of C_t and A_t and, given these, we solve the durable choice problem. The first order conditions for C_t and A_t yield a standard Euler equation:

$$\lambda_t = \beta \frac{(1+i_t)}{\Pi_{t+1}} \lambda_{t+1}, \tag{3.43}$$

$$\lambda_t = (C_t)^{-\sigma}, \tag{3.44}$$

where λ_t is the Lagrange multiplier on the aggregate budget constraint and $\Pi_t = P_t/P_{t-1}$ is the gross inflation rate for nondurable goods. The optimal choice for $D_{i,t}$ conditional on household i making an adjustment at time t solves:

$$\begin{aligned} \max_{D_{i,t}} E_t \sum_{s=0}^{\infty} (\beta\theta^d)^s \left[\psi_{t+s}\varepsilon_{\psi t+s} \frac{\left((1-(1-\chi)\delta)^s D_{i,t}\right)^{1-\sigma_d}}{1-\sigma_d} - \lambda_{t+s}\iota(1-(1-\chi)\delta)^s D_{i,t} \right] \\ - \lambda_t p_t^x D_{i,t} + \sum_{s=0}^{\infty} \beta^s \theta^{d^{s-1}} (1-\theta^d)(1-(1-\chi)\delta)^s \lambda_{t+s} p_{t+s}^x D_{i,t}, \end{aligned} \tag{3.45}$$

where θ^{d^s} is the probability that the durable stock chosen at t survives s periods into the future and $p_t^x = P_t^x/P_t$ is the relative price of the durable good. The first term in brackets, $\psi_t\varepsilon_{\psi t}\frac{\left((1-(1-\chi)\delta)^s D_{i,t}\right)^{1-\sigma_d}}{1-\sigma_d}$, captures the utility flow of the durable stock in $t+s$. The term $\lambda_{t+s}\iota(1-(1-\chi)\delta)^s D_{i,t}$ measures the cost of operating the durable stock in utils of time $t+s$ while the term $p_t^x D_{i,t}$ measures the cost of purchasing the durable stock in period t. The last term captures the resale value of the durable stock if an adjustment opportunity arrives in period $t+s$. All households that re-optimize at t solve the same problem, therefore, we can denote their optimal durable stock choice by D_t^*. It can be shown that D_t^* satisfies the following Calvo-type specification[13]:

[13]This formulation is analogous to the optimal re-set price in a standard model with Calvo-type price rigidities.

$$D_t^* = \left(\frac{\Omega_{1,t}}{\Omega_{2,t}}\right)^{\frac{1}{\sigma_d}}, \tag{3.46}$$

$$\Omega_{1,t} = \psi_t \varepsilon_{\psi t+s} + \beta\theta^d (1-(1-\chi)\delta)^{1-\sigma_d}\Omega_{1,t+1}, \tag{3.47}$$

$$\Omega_{2,t} = \lambda_t \left(p_t^x + \iota - \frac{(1-(1-\chi)\delta)}{(1+i_t)\Pi_{t+1}^{-1}} p_t^x (cg_{t+1}-1) \right) + \beta\theta^d (1-(1-\chi)\delta)\Omega_{2,t+1}, \tag{3.48}$$

where $cg_{t+1} = 1 + \frac{p_{t+1}^x}{p_t^x}$ is the capital gain from selling the stock of durable goods in period $t+1$. The optimal aggregate purchases of durable goods are given by:

$$X_t = (1-\theta^d)\left[D_t^* - (1-(1-\chi)\delta)D_{t-1}\right] + \chi\delta D_{t-1}, \tag{3.49}$$

The first term captures new durables purchases by households that are allowed to reset their durable stock and the second term captures maintenance expenditures which are incurred by all households, regardless of their re-set status. Aggregate nondurable expenditure, ND_t, is defined as the sum of consumption of the nondurable good plus operating costs:

$$ND_t = C_t + \iota D_t, \tag{3.50}$$

3.6.2 Regional Demand

The composite nondurable good is given by:

$$ND_t = \left[\gamma_c ND_{1t}^{\frac{\eta_c-1}{\eta_c}} + (1-\gamma_c) ND_{2t}^{\frac{\eta_c-1}{\eta_c}}\right]^{\frac{\eta_c}{\eta_c-1}}, \tag{3.51}$$

where ND_{rt} denotes consumption of nondurable goods produced in region r.

The composite X_t aggregates durable goods produced in each region according to:

$$X_t = \left[\gamma_x X_{1t}^{\frac{\eta_x-1}{\eta_x}} + (1-\gamma_x) X_{2t}^{\frac{\eta_x-1}{\eta_x}}\right]^{\frac{\eta_x}{\eta_x-1}}, \tag{3.52}$$

where X_{rt} denotes purchases of durable goods produced in region r.

Similarly, total labor supply is a CES aggregate over labor supplied to each region:

$$L_t = \left[\gamma_\ell L_{1t}^{\frac{\eta_\ell-1}{\eta_\ell}} + (1-\gamma_\ell) L_{2t}^{\frac{\eta_\ell-1}{\eta_\ell}}\right]^{\frac{\eta_\ell}{\eta_\ell-1}}, \tag{3.53}$$

where L_{rt} denotes hours worked in region r. η_ℓ governs the elasticity of substitution of labor supply across regions and γ_ℓ captures the relative preference of households to work in region 1.

Given these preferences, we have the following regional demand schedules:

$$ND_{1t} = \gamma_c ND_t \left(\frac{P_{1t}}{P_t}\right)^{-\eta_c} \qquad X_{1t} = \gamma_x X_t \left(\frac{P^x_{1t}}{P^x_t}\right)^{-\eta_x} \tag{3.54}$$

$$ND_{2t} = (1-\gamma_c) ND_t \left(\frac{P_{2t}}{P_t}\right)^{-\eta_c} \qquad X_{2t} = (1-\gamma_x) X_t \left(\frac{P^x_{2t}}{P^x_t}\right)^{-\eta_x}. \tag{3.55}$$

With the corresponding CES price indices given by:

$$P_t = \left[\gamma_c P_{1t}^{1-\eta_c} + (1-\gamma_c) P_{2t}^{1-\eta_c}\right]^{1-\eta_c} \tag{3.56}$$

$$P^x_t = \left[\gamma_x P^{x\ 1-\eta_x}_{1t} + (1-\gamma_x) P^{x\ 1-\eta_x}_{2t}\right]^{1-\eta_x}. \tag{3.57}$$

3.6.3 Nondurable Goods Production

Nondurable goods are produced by perfectly competitive firms using a technology linear in labor. The problem of a representative firm located in region r is:

$$\max \quad \pi_t = P_{rt} Y_{rt} - W_{rt} L_{rt} \qquad \text{s.t.} \quad Y_{rt} = Z_t Z_r L_{rt}, \tag{3.58}$$

where Z_t is aggregate TFP, Z_r is a region-specific productivity shifter, and W_{rt} is the nominal wage prevailing in region r. Under flexible prices, profit-maximizing firms set prices equal to nominal marginal costs:

$$P_{rt} = \frac{W_{rt}}{Z_t}. \tag{3.59}$$

3.6.4 Durable Goods

In each region, a continuum of perfectly competitive firms operate with a technology that transforms local nondurable goods into durable goods subject to a congestion externality (McKay and Wieland, 2022). This externality introduces decreasing returns to scale in durable goods production. The technology of a representative firm i in region r is:

$$X_{irt} = Z_r^x M_{irt} \left(\frac{X_{rt}}{\bar{X}_r} \right)^{-\zeta}, \tag{3.60}$$

where Z_r^x is region-specific productivity in durable production, M_{irt} are inputs of locally-sourced nondurable goods, X_{rt} is total durable good production in region r and $\bar{X}_r$ is durable production of r on the balanced growth path. The term $\left(\frac{X_r}{\bar{X}_r}\right)^{-\zeta}$ captures the congestion externality.

The optimal price of the perfectly competitive firm is:

$$P_{rt}^x = \frac{P_{rt}}{Z_r^x} \left(\frac{X_{rt}}{\bar{X}_r} \right)^{\zeta}. \tag{3.61}$$

Firms set prices equal to region-specific marginal costs. P_r^x is the price of non-durable goods produced in region r. Parameter ζ, the inverse price-elasticity of durable supply, governs the strength of the congestion externality.

3.6.5 Aggregate TFP growth

Total factor productivity Z_t grows at a constant rate $g_{zt} = \bar{g}$.

3.6.6 Labor Unions

There is a continuum of labor unions in each region. Each union j supplies a differentiated labor type to a local competitive firm that aggregates all labor types and sells them at price W_{rt} to local nondurable producers. The aggregator is,

$$L_{rt}^{d} = \left[\int_0^1 L_{rt}^{d}(j)^{\frac{\varepsilon^w - 1}{\varepsilon^w}} dj \right]^{\frac{\varepsilon^w}{\varepsilon^w - 1}}, \tag{3.62}$$

where ε^w measures the elasticity of substitution across labor types within a region and L_{rt}^d is labor demand from nondurable producers located in region r.

The labor demand faced by union j is given by:

$$L_{rt}^{d}(j) = L_{rt}^{d} \left(\frac{W_{rt}(j)}{W_{rt}} \right)^{-\varepsilon^w}. \tag{3.63}$$

The union chooses wages to maximize the average expected utility across households subject to a Calvo-type friction. The problem of a union j in region r that gets to re-optimize its wage at time t is as follows:

$$\max_{W_{rt}^{*}(j)} \sum_{s=0}^{\infty} (\beta\theta^w)^s \left[U_{c_{t+s}} \frac{W_{rt+s|t}^{*}(j)}{P_{t+s}} L_{rt+s|t}^{d}(j) - U_{L_{r,t+s}} L_{rt+s|t}^{d}(j) \right]$$

where:

$$L_{rt+s|t}^{d}(j) = L_{rt+s}^{d} W_{rt+s}^{\varepsilon^w} W_{rt}^{*}(j)^{-\varepsilon^w}, \tag{3.64}$$

$$U_{L_{rt+s}} = \varphi_{t+s} L_{t+s}^{\nu} \gamma_{\ell}^{\frac{1}{\eta_\ell}} L_{rt+s}^{-\frac{1}{\eta_\ell}}, \tag{3.65}$$

$$U_{c_{t+s}} = \lambda_{t+s}. \tag{3.66}$$

θ^{ws} is the probability that the wage contract survives s periods into the future, U_{ct} is the marginal utility of consumption and $U_{L_{rt}}$ is the marginal disutility of labor supplied to region r. In absence of an adjustment opportunity, the prevailing wage in $t+s$ stays fixed at its period t value. The labor union chooses $W^*_{rt}(j)$ such that it maximizes household utility over the expected duration of the wage contract. Since the problem of all unions adjusting in period t in region r is symmetric, we omit indexing the optimal reset wage by j. The optimal wage contract solves:

$$w^*_{rt} = \frac{\varepsilon^w}{\varepsilon^w - 1} \frac{F_{1rt}}{F_{2rt}} \tag{3.67}$$

$$F_{1rt} = \gamma_\ell^{-\frac{1}{\eta_\ell}} \varphi_t L_t^{\nu - \frac{1}{\eta_\ell}} L_{rt}^{\frac{1}{\eta_\ell}} L^d_{rt} w_{rt}^{\varepsilon^w} + \beta\theta^w \Pi_{t+1}^{\varepsilon^w} F_{1r,t+1} \tag{3.68}$$

$$F_{2rt} = \lambda_t L^d_{rt} w_{rt}^{\varepsilon^w} + \beta\theta^w \Pi_{t+1}^{\varepsilon^w - 1} F_{2r,t+1} \tag{3.69}$$

where lowercase letters denote real wages and Π_t denotes the gross price inflation rate. With perfectly flexible wages, optimality implies that wages equal a constant mark-up over the marginal rate of substitution. Away from this scenario, the wage mark-up will fluctuate across time and unions. This wage dispersion induces a wedge (in terms of *utils*) between hours worked, L_{rt}, and effective hours worked, L^d_{rt}:

$$L_{rt} = s^w_{rt} L^d_{rt}, \tag{3.70}$$

where s^w_{rt} measures the degree of wage dispersion across unions within a region. This wedge evolves according to:

$$\begin{aligned} s^w_{rt} &= \int_0^1 \left(\frac{W_{rt}(j)}{W_{rt}} \right)^{-\varepsilon^w} dj \\ &= (1-\theta^w) \left(\frac{W^*_{rt}}{W_{rt}} \right)^{-\varepsilon^w} + \theta^w \left(\frac{W_{rt-1}}{W_{rt}} \right)^{-\varepsilon^w} s^w_{rt-1} \\ &= (1-\theta^w) \left(\frac{w^*_{rt}}{w_{rt}} \right)^{-\varepsilon^w} + \theta^w \left(\frac{w_{rt-1}}{w_{rt}} \Pi_t \right)^{-\varepsilon^w} s^w_{rt-1}. \end{aligned} \tag{3.71}$$

The real wage index for jobs located in region r satisfies:

$$w_{rt} = \left[(1-\theta^w) {w^*_{rt}}^{1-\varepsilon^w} + \theta^w \left(w_{rt-1} \Pi_t^{-1} \right)^{1-\varepsilon^w} \right]^{\frac{1}{1-\varepsilon^W}}. \tag{3.72}$$

Lastly, the aggregate real wage, w_t, is characterized by the following CES index:

$$w_t = \left[\gamma_\ell w_{1t}^{1-\eta_\ell} + (1-\gamma_\ell) w_{2t}^{1-\eta_\ell} \right]^{\frac{1}{1-\eta_\ell}}. \tag{3.73}$$

3.6.7 Monetary Policy

The central bank sets the real interest rate according to the following monetary policy rule:

$$1 + r_t = (1 + r_{ss}) \left(\frac{L_t}{L} \right)^{\phi_L} \varepsilon_{rt}, \tag{3.74}$$

where $\bar{r}$ is the real rate along the balanced growth path, L is aggregate labor along the balanced growth path and ε_{rt} is an exogenous shock to the real interest rate.

3.6.8 Market Clearing

Regional goods market clearing implies:

$$Y_{rt} = C_{rt} + \frac{P^x_{rt}}{P_{rt}} X_{rt} \qquad \forall \quad r = 1, 2. \tag{3.75}$$

Aggregate output, Y_t, is given by:

$$Y_t = \frac{P_{1t}}{P_t} Y_{1t} + \frac{P_{2t}}{P_t} Y_{2t}. \tag{3.76}$$

Lastly, asset market clearing implies:

$$A_t = 0. \tag{3.77}$$

3.6.9 Stationary Equilibrium

We normalize all prices by the nondurable price index, P_t, and make the economy stationary by expressing all the growing variables relative to aggregate productivity. We denote with lowercase letters these relative variables and solve the model around a zero-inflation balanced growth path with $y_{ss} = l_{ss} = 1$, $w_{ss} = w_{i,ss} = 1$, and equal-sized regions, $p_{1,ss}y_{1,ss} = p_{2,ss}y_{2,ss} = .5y_{ss}$. The stationary equilibrium is defined by a sequence of productivity-adjusted quantities and wages, a sequence of goods prices $\{p_{it}, p^x_{it}\}$, plus a sequence for the real rate $\{r_t\}$ such that, given the realizations for $\{\varepsilon_\psi, \varepsilon_r\}$: (1) household utility and firm profits are maximized, (2) durable and nondurable goods markets clear, and (3) the asset market clears.

3.6.10 Calibration

Table 3.2 shows the calibrated parameters. We calibrate the model to a quarterly frequency. We set the parameters for the household sector as follows. For the degree of household inattention we choose $\theta = .93$, in line with the estimates of Auclert et al. (2020). We calibrate the effective discount factor to $\frac{\beta}{1+g_z} = .99$ to match the average real rate during our sample. We calibrate the depreciation rate to match the average depreciation of the stock of private fixed assets plus durable goods in the BEA fixed asset table between 1947-2019. Operating and maintenance costs are computed as in McKay and Wieland (2022). Operating costs include expenditures in households utilities (PCE), expenditures in motor vehicle fuels and fluids (PCE), and taxes on the housing sector. We calibrate $\iota = .005$ to match the average operating costs as a share of the durable goods stock. Maintenance costs are calculated by aggregating household and motor vehicle maintenance expenditures from the Personal Consumption Expenditures (PCE) data, as well as intermediate goods and services from the housing output table. We set $\chi = .17$ to

Table 3.2. Calibrated Parameters.

Parameters		Value
$1/\sigma$	Nondurable EIS	.25
$1/\sigma^d$	Durable EIS	.25
$\bar{\psi}$	Taste for Durables	6.3
ζ	Inverse Price Elasticity of Durable Supply	.1
ι	Operating cost	0.005
χ	Maintenance cost	0.17
δ	Depreciation rate	0.017
$\frac{\beta}{(1+g_z)}$	Effective Discount factor	0.99
$\bar{g}$	Productivity Growth	.005
Y/Z	Output to TFP Ratio	1
ν	Inverse Frisch elasticity	1
ϕ_L	Taylor Rule Coefficient - Output	0
η_c	Elasticity of Substitution - Nondurables	5
η_x	Elasticity of Substitution - Durables	7
η_ℓ	Elasticity of Substitution - Labor Supply	5
ε_w	Elasticity of Substitution - Labor Demand	7
θ	Degree of Information Stickiness	.93
θ^d	Calvo Parameter - Durable Consumption	.8
θ^w	Degree of Wage Stickiness	.9
γ_x	Regional Bias in Durable Demand	.5
γ_ℓ	Regional Bias in Labor Supply	.5
γ_c	Regional Bias in Nondurable Demand	.5
Z_1^x	Productivity of Investment Sector in Region 1	1.04
Z_2^x	Productivity of Investment Sector in Region 2	.95
Z_1	Productivity of Nondurable Sector in Region 1	.98
Z_2	Productivity of Nondurable Sector in Region 2	1.01

match the average ratio of maintenance costs to depreciation. The taste for durable goods, $\bar{\psi}$, is calibrated to match the average ratio of nondurable consumption to the durable stock during our sample. We set θ^d,the Calvo friction in the durable choice problem, at $\theta^d = .8$. This implies an elasticity of investment to the real rate of -4 (Gilchrist and Himmelberg (1995), Auclert and Rognlie (2018)). We set the elasticity of substitution for durables equal to $\eta_x = 7$ and the elasticity of substitution for nondurables to $\eta_c = 5$. Higher values for these elasticities imply stronger expenditure switching effects in response to changes in prices across regions. We choose these values to reflect the relatively lower tradability of nondurable goods which we characterize as mostly being labor services. Lastly, we set the elasticity of substitution of labor supply between regions at $\eta_\ell = 5$.

We calibrate the production side as follows. We normalize the output-to-TFP ratio to one in the stationary equilibrium. The deterministic component of TFP growth, g_Z, is calibrated to match the average quarterly growth rate of real GDP per capita over the period 1947-2019. We calibrate the inverse supply elasticity of durable goods to $\zeta = .1$ based on House and Shapiro (2008). We set the elasticity of substitution of labor demand across unions to $\varepsilon_w = 7$. This is in the middle range of values used in the literature and implies a wage markup of 15% in the steady state. We set the degree of wage stickiness to $\theta_w = .9$, which implies that unions re-optimize wages on average every 10 quarters.

We set the productivity parameters in the durables sector, Z_1^x and Z_2^x, and nondurables sector, Z_1 and Z_2, to generate heterogeneity in the industrial composition between regions. Specifically, we choose Z_1^x such that the participation of region 1 in total durable production is 55%. Then Z_2^x is set to guarantee that the relative aggregate price of durable goods is equal to one in steady state (that is, $p_{ss}^x = 1$). Lastly, we choose productivity values in the nondurable sector to ensure that both regions have the same level of steady-state output. As a result, Region 1 is relatively specialized in the production of durable goods, while Region 2 is relatively specialized in nondurable goods.

3.6.11 Matching the Data

We model boom-bust recessions as the result of a fake news shock to households' taste for durables ψ_t. We assume a two-wave information structure to construct the fake news shock which we briefly describe next. Appendix C.4 provides further details. In period $t = 0$, news arrive about an increase in ψ_t beginning in period $s > 0$. This constitutes the first wave of information. Starting at $t = 0$, agents gradually learn of this future path at a rate $1 - \theta$, where $\theta \in [0, 1)$ captures the degree of information stickiness. Households who update their information set expect a higher return on durable consumption beginning in period s. In anticipation of this, they build up their durable stock, which boosts output and the investment share. Because households update their information set infrequently, this results in a gradual build-up of the aggregate durable stock. In period $z < s$, a second wave of information begins. In each period $t \geq z$, a share $1 - \theta$ of agents who expect ψ_t to increase at time s learn that the taste for durables will not change. Households who receive this new piece of information find themselves with an excess stock of durable goods they wish to liquidate. Initially, this results in a deceleration of the process of accumulation of durable goods at the aggregate level. Accumulation decelerates but does not come to an immediate halt. This is because the stock of households that still expect ψ_t to increase is large relative to those who learn of the second wave. As the second wave of information disseminates, the liquidation of excess stocks generates a bust in output, a sustained decrease in the investment share, and a slow recovery of the economy relative to the elevated pretrend.

We generate a V-shaped recession pattern for the pre-1990 and Covid recessions using a temporary monetary policy shock. In practice, any transitory shock in the model will generate a V-shaped recession in the model. We choose monetary policy shocks because they are considered important drivers of pre-1990 recessions (Coibion, 2012; Romer and Romer, 2023). One can interpret the Covid recession as a period in which the price of consumption today relative to the future—the real interest rate—was high due to health risk associated with certain types of consumption. An unexpected increase in the real interest rate induces intertemporal substitution away from consumption and investment that cause output to temporarily drop below trend before bouncing back once the monetary shock recedes. Due to the importance of intertemporal motives in such shocks, we call them intertemporal shocks.

We model the relatively weaker trend of the more exposed region as the result of a change in γ_x, the regional bias in durable demand. A change in γ_x triggers the reallocation of durable household purchases across regions, holding total purchases fixed. This allows us to compute the path of γ_x that rationalizes a given time path for the differential in growth rates between regions, absent any national recessions. Given the linearity of the model, we can then additively feed in the aggregate shocks that generate national recessions, as discussed above.

We solve the model in the sequence space using the methods developed by Auclert et al. (2021). We then simulate the path of the economy for different shocks that generate heterogeneous recessions and recovery dynamics. Importantly, we keep all the parameters that affect the steady state fixed across recessions, which means that all differential dynamics can be traced back to the type of shock that hits the economy. Next, using the simulated data, we compute the model analog of the relative recession impulse response function in Figure 3.4. Specifically, we estimate the same regression (3.33) using a Bartik instrument in the model as we do in the data.

We estimate the path of γ_x[14] and the parameters that govern the shock processes to match the impulse responses in Figure 3.4. We assume AR(1) processes for both the monetary policy and taste for durables shocks:

$$\varepsilon_t^r = \rho_r e_{t-1}^r + e_t^r, \quad e_t^r \overset{\text{iid}}{\sim} N(0, \sigma_r^2) \tag{3.78}$$

$$\varepsilon_t^\psi = \rho_\psi \varepsilon_{t-1}^\psi + e_t^\psi, \quad e_t^\psi \overset{\text{iid}}{\sim} N(0, \sigma_\psi^2), \tag{3.79}$$

where ρ_x and σ_x are the persistence and standard deviation of shock x. For each alternative explanation, we estimate the persistence of the shock processes. We also estimate the timing of each information wave, s and z. Let $f(\mathbf{\Gamma})$ denote the vector of model-implied cross-sectional impulse responses evaluated at parameter values $\mathbf{\Gamma}$. Let $\boldsymbol{\beta}$ denote the vector of empirical impulse responses estimated in Section 3.5 .

We estimate the parameters by solving:

$$\hat{\mathbf{\Gamma}} = argmin_{\mathbf{\Gamma}} \quad (\boldsymbol{\beta} - f(\mathbf{\Gamma}))' \mathcal{W}^{-1} (\boldsymbol{\beta} - f(\mathbf{\Gamma})), \tag{3.80}$$

where $\mathcal{W}$ is a diagonal weighting matrix containing the estimated variances of the elements of β and $\hat{\mathbf{\Gamma}}$ is the value of the parameter vector that minimizes the weighted distance between the

[14]In particular, we estimate a single path of γ_x, which means that both types of recessions occur around the same relatively weaker pretrends.

Table 3.3. Estimated Parameters.

		Value
Intertemporal *vs.* Boom-bust		
Persistence of Monetary Policy Shock	ρ_r	.946
Persistence of Taste Shock	ρ_ψ	1
Fake News Date	s	17
Second Wave of News Date	z	14

model-implied and empirical impulse responses. Table 3.3 presents the estimated parameters. Throughout the exercise, we calibrate the standard deviation of each aggregate shock to match the average depth of the recession in our sample, which is 5.3%.

In Figure 3.8 we show that when recessions are driven by different aggregate shocks, the model can generate impulse response functions that closely match the cross-sectional estimates for the 1990-1, 2001, and 2007-9 recessions. Regions that are more exposed to the boom-bust recession have a relatively higher above-trend growth rate due to the prior boom. Recoveries in these regions will look weaker because they can never catch up to this inflated pre-recession trend. The estimation favors a slight downward trend in the more exposed regions due to the negative pretrend prior to the boom.

Intertemporal shocks, such as a monetary policy shock, do not affect the pretrend and represent only a temporary departure from it. Thus, these shocks predict a V-shaped relative recovery in the absence of pretrends. However, our estimates point to a significantly lower trend growth rate in the more exposed States which helps the model match the observed recovery in the data.

3.6.12 Policy Analysis

We next show that distinguishing between boom-bust cycles and intertemporal shocks is important for policy in the recovery. Specifically, for each recession we solve for the path of the real interest rate R^* that closes the aggregate output gap starting at the recession trough. We

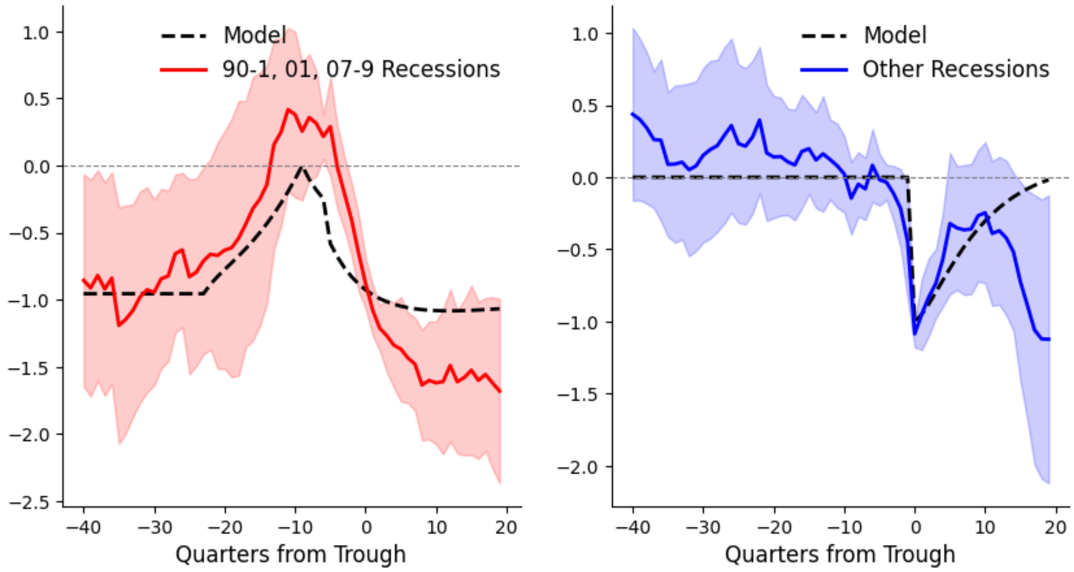


Figure 3.8. Matching the Recession IRF - Liquidationist *vs.* Intertemporal.

scale each recession to a 5.2% peak-to-trough decline in aggregate output, which is the empirical average for both types of recessions. Figure 3.9 shows the 10-year R^* zero coupon rate computed using the expectations hypothesis.

The model implies that long-run R^* persistently declines following a boom-bust cycle, reaching a trough of 20 basis points. In contrast, the V-shaped monetary policy recession predicts no change in long-run R^*.

These distinct patterns for R^* are due to the dynamics of the capital stock. Following a boom-bust cycle, the economy finds itself with excess capital. Investment is persistently low and aggregate demand needs to be supported using a low real interest rate. In contrast, the intertemporal recession considered here lasts only one period, so closing the output gap from the recession trough onward amounts to offsetting the contractionary monetary shock. More generally, if an intertemporal recession lasted for several periods and policy closed output gaps only after the recession had ended, delayed investment would leave the capital stock below steady state. In that case, pent-up investment demand would push R^* up during the recovery. This implies that the degree of monetary accommodation required in the recovery depends importantly on the type of shock.

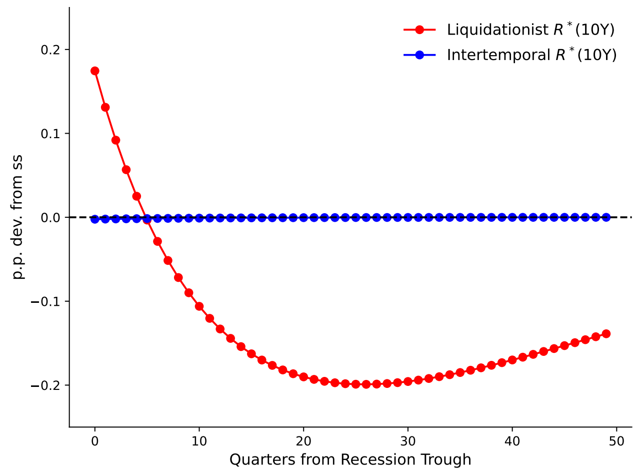


Figure 3.9. R_t^* - Liquidationist *vs.* Intertemporal.

3.7 Conclusion

The recoveries from the 1990-1, 2001, and 2007-9 recessions have been weak relative to other postwar business cycles. We provide new cross-State evidence that suggests the recovery patterns are due to different shocks: Boom-bust cycles in the 1990-1, 2001, and 2007-9 recessions give rise to weak recoveries and intertemporal shocks in the pre-1990 and 2020 recessions predict strong recoveries. Our model that quantitatively matches these patterns shows that recessions caused by boom-bust cycles require extended periods of accommodative policy relative to recessions caused by intertemporal shocks.

Acknowledgements

Chapter 3, in full, is based on the paper from Paula Donaldson Paula and Johannes Wieland, "Why are Some Recoveries Weak and Others Strong?" currently being prepared for submission. The dissertation author is one of the primary authors of this paper.

Appendix A

Appendix to Chapter 1

A.1 Decomposition Framework - Additional Derivations and Results

A.1.1 Static Setting - Further Details

This appendix presents the derivations for a generalized static setting with two-way feedback between aggregate variables. In the following, I reproduce the system of equations that characterize a static economic:

$$
\begin{aligned}
G_t &= \alpha R_t + u_{gt} \\
R_t &= \delta G_t + u_{rt} \\
u_{kt} &\sim \mathrm{N}(0, \sigma_k^2) \quad \forall\, k = g, r
\end{aligned}
\qquad
\begin{aligned}
Y_{it} &= \beta \times s_{ig} G_t + \gamma \times s_{ir} R_t + u_i + u_t + u_{it}^y \\
s_{ig} &\sim \mathrm{N}(\bar{s}, \sigma_{s_g}^2) \\
s_{ir} &= \phi s_{ig} + u_i^{s_r} \\
u_i^{s_r} &\sim N(0, \sigma_{s_r}^2), \quad u_{it}^y \sim N(0, \sigma_y^2).
\end{aligned}
$$

With two-way feedback, the process for the aggregate variable G and R can be rewritten as:

$$G_t = \frac{1}{1-\delta\alpha} u_{gt} + \frac{\alpha}{1-\delta\alpha} u_{rt}, \tag{A.1}$$

$$R_t = \frac{\delta}{1-\delta\alpha} u_{gt} + \frac{1}{1-\delta\alpha} u_{rt}. \tag{A.2}$$

Next, I outline the general formulas for each step of the decomposition.

Cross-sectional Step

The reduced-form regression from the cross-sectional step is given by:

$$Y_{it} = b s_{ig} \varepsilon_{gt} + c s_{ig} \varepsilon_{rt} + \lambda_i + \lambda_t + e_{yt}. \tag{A.3}$$

The regression coefficients $\hat{b}$ and $\hat{c}$ converge to:

$$\mathbb{E}[\hat{b}] = \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta}, \tag{A.4}$$
$$\mathbb{E}[\hat{c}] = \frac{\beta\alpha}{1-\alpha\delta} + \frac{\tilde{\gamma}}{1-\alpha\delta}. \tag{A.5}$$

Note that, due to the presence of two-way feedback, we cannot decompose $\hat{b}$ into two additive terms as before. The first term on the right-hand side is a rescaled version of the portable elasticity where the rescaling captures the contemporaneous feedback between aggregate variables. Similarly for the HEDGE term. However, this is not an impediment to apply the decomposition as I show next.

Time Series Step

The estimated responses of R_t to each aggregate shock are given by:

$$R_t = v\varepsilon_{gt} + a\varepsilon_{rt} + e_{rt}, \qquad \mathbb{E}[\hat{v}] = \frac{\delta}{1-\delta\alpha}, \qquad \mathbb{E}[\hat{a}] = \frac{1}{1-\delta\alpha}. \tag{A.6}$$

The hypothetical ε_{rt} shock - denoted by $\tilde{\varepsilon}_{rt}$ - satisfies the following:

$$\tilde{\varepsilon}_{rt} = \frac{\hat{v}}{\hat{a}} = \delta. \tag{A.7}$$

Last Step

As pointed out above, with two-way feedback we cannot in general neatly decompose the TWFE estimate into two additive terms but we can still consistently estimate the portable elasticity in the same way as before. Evaluating the effect of $\tilde{\varepsilon}_{rt}$ using $\hat{c}$ from the cross-sectional step and subtracting it from $\hat{b}$ yields:

$$\begin{aligned}\hat{\beta}^P &= \hat{b} - \hat{c} \times \tilde{\varepsilon}_{rt} \\ &= \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta} - \frac{\beta\alpha\delta}{1-\alpha\delta} - \frac{\tilde{\gamma}\delta}{1-\alpha\delta} \\ &= \beta.\end{aligned} \tag{A.8}$$

Monte Carlo Simulations with Two-Way Feedback

The performance of the decomposition is illustrated using Monte Carlo simulations. I simulate data from an economy characterized by the system of equations presented above with $\alpha \neq 0$ and estimate three different specifications: (i) Baseline, (ii) Decomposition, and (iii) Control Function. Baseline corresponds to the two-way fixed-effects regression. Decomposition and control function correspond to the results of applying the framework presented above or the control function approach. Figure A.1 presents the distribution of the estimated coefficients for each strategy in an economy with $\beta = \gamma = \delta = \phi = \alpha = .5$. The population value for the portable effect is .5 and the omitted variable bias due to HEDGE is .33. There are two takeaways. First, the baseline estimates are biased and centered on $\beta^P + \Omega = .83$. Second, both the decomposition and control function approach consistently estimate β^P.

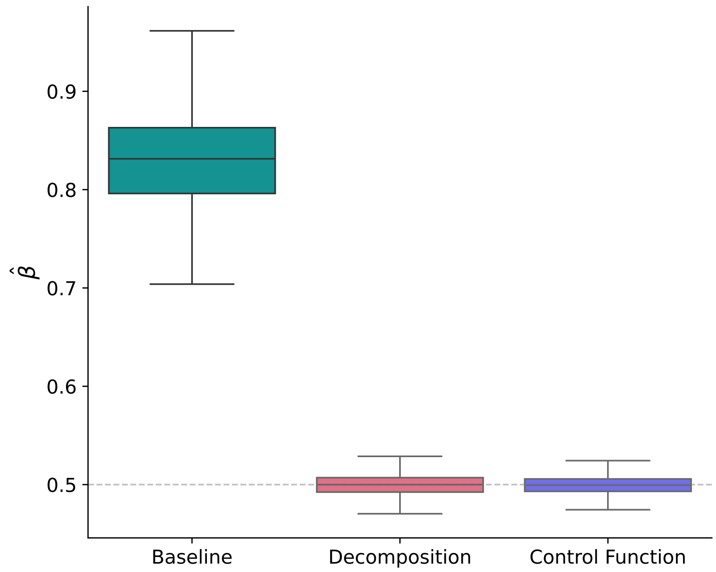


Figure A.1. Static Monte Carlo Simulations with Two-Way Feedback.

Note: Distribution of point estimates for the *Baseline*, *Decomposition Framework* and *Control Function* estimation strategies. Based on 1000 repetitions with sample size of $n = 100$ and $t = 300$ and parameters set to $\delta = \beta = \gamma = \phi = \alpha = .5$.

A.1.2 Decomposition a la Sims and Zha (2006) - Further Details

This appendix presents derivations for the *ex-post* decomposition in a more general economy than the AR(1) example in the main text.

Let the data-generating process be characterized by

$$\begin{aligned}
s_{ig} &= u_i^{s_g}, \qquad s_{ir} = \phi s_{ig} + u_i^{s_r}, \\
G_t &= \sum_{j=1}^{J} \rho_{gj} G_{t-j} + u_{gt}, \\
R_t &= \sum_{m=1}^{M} \rho_{rm} R_{t-m}, + \sum_{n=0}^{N} \delta_n G_{t-n} + u_{rt}, \\
Y_{it+h} &= \sum_{s=0}^{S} \beta_s s_{ig} G_{t+h-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t+h-k} + \sum_{l=1}^{L} \psi_l Y_{i,t+h-l} + u_{it+h}^{y} \quad \forall h \geq 0, \\
u_{kt} &\sim \mathrm{N}(0, \sigma_k^2),\ u_i^{s_k} \sim \mathrm{N}(0, \sigma_{s_k}^2)\ \forall\ k = g, r, \quad u_{it}^{y} \sim \mathrm{N}(0, \sigma_y^2).
\end{aligned} \tag{A.9}$$

The derivations that follow set $\psi_l = 0$ for notational tractability. This is without loss of generality.

The total, portable and HEDGE effects of a shock to G_t in period $t+h$ are given by:

$$\beta_h^T = \underbrace{\sum_{s=0}^{S} \beta_s \frac{\partial G_{t+h-s}}{\partial u_{gt}}}_{\beta_h^P} + \underbrace{\sum_{k=0}^{K} \gamma_k \phi \frac{\partial R_{t+h-k}}{\partial u_{gt}}}_{\Omega_h}. \tag{A.10}$$

where the effects are evaluated at $s_{ig} = 1$. The total derivative of the cross-sectional outcome at horizon h with respect to u_{rt} is given by:

$$\gamma_h^T = \sum_{k=0}^{K} \gamma_k \phi \frac{\partial R_{t+h-k}}{\partial u_{rt}}. \tag{A.11}$$

Reduced Form Analysis

Consider the following panel local projections with two-way fixed effects:

$$Y_{it+h} = b_h \times s_{ig}\varepsilon_{gt} + c_h \times s_{ig}\varepsilon_{rt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{A.12}$$

where ε_{gt} and ε_{rt} are the innovations to G and R, respectively. The regression coefficient $\hat{b}_h$ converges to:

$$\mathbb{E}[\hat{b}_h] = \sum_{s=0}^{S} \beta_s \frac{cov[G_{t+h-s}, \varepsilon_{gt}]}{v[\varepsilon_{gt}]} + \sum_{k=0}^{K} \gamma_k \phi \frac{cov[R_{t+h-k}, \varepsilon_{gt}]}{v[\varepsilon_{gt}]}, \tag{A.13}$$

$\forall\, s \in S$ such that $h - s \geq 0$ and $\forall\, k \in K$ such that $h - k \geq 0$. The regression coefficient $\hat{c}_h$ converges to:

$$\mathbb{E}[\hat{c}_h] = \sum_{k=0}^{K} \gamma_k \phi \frac{cov[R_{t+h-k}, \varepsilon_{rt}]}{v[\varepsilon_{rt}]}, \tag{A.14}$$

$\forall\, k \in K$ such that $h - k \geq 0$. If $\frac{cov[R_{t+h}, \varepsilon_{rt}]}{v[\varepsilon_{rt}]} = \frac{cov[R_{t+h}, \varepsilon_{gt}]}{v[\varepsilon_{gt}]}$ for all $h \in H$, then the cross-sectional step directly provides us with an estimate of Ω_h at each horizon. However, whenever the conditional path of R given each aggregate shock differs, an additional correction is needed. Implementing this correction requires information from time-series analysis. Basically, we need estimates for the path of R conditional on the realization of each aggregate shock.

The time-series step estimates the path of R conditional on each aggregate shock using:

$$R_{t+h} = v_h \varepsilon_{gt} + a_h \varepsilon_{rt} + e_{rt+h}. \tag{A.15}$$

The coefficients v_h and a_h converge to

$$\mathbb{E}[\hat{a}_h] = \frac{cov[R_{t+h}, \varepsilon_{rt}]}{v[\varepsilon_{rt}]} = \sum_{m=1}^{M} \rho_m \frac{cov[R_{t+h-m}, \varepsilon_{rt}]}{v[\varepsilon_{rt}]} + \frac{cov[u_{rt+h}, \varepsilon_{rt}]}{v[\varepsilon_{rt}]}, \tag{A.16}$$

$$\mathbb{E}[\hat{v}_h] = \frac{cov[R_{t+h}, \varepsilon_{gt}]}{v[\varepsilon_{gt}]} = \sum_{n=0}^{N} \beta_{rg,n} \frac{cov[R_{t+h-n}, \varepsilon_{gt}]}{v[\varepsilon_{gt}]}, \tag{A.17}$$

for all $m \in M$ such that $h - m \geq 0$ and for all $n \in N$ such that $h - n \geq 0$. Following the notation in the body of the paper, I denote with $\boldsymbol{R}_{g,h}$ and $\boldsymbol{R}_{r,h}$ the impulse response of R to each innovation at horizon h.

The coefficients $\hat{c}_h$ from the cross-sectional step identify the differential response of the unit level outcome to a unit increase in ε_{rt}. This ε_{rt} shock implies a path for R given by $\hat{a}_h$ for each h. However, to construct the HEDGE term we need the path for R given by the coefficients $\hat{v}_h$. Following Sims and Zha (2006), I outline how to replicate $\boldsymbol{R}_g$ using a series of contemporaneous innovations to R dated $s \geq 0$ for all $s \in H$.

For $h = 0$, we want to find the size of the ε_{r0} shock that equalizes both paths on impact:

$$\boldsymbol{R}_{g,0} = \boldsymbol{R}_{r,0}\tilde{\varepsilon}_{r0} \qquad \Rightarrow \qquad \tilde{\varepsilon}_{r0} = \frac{\boldsymbol{R}_{g,0}}{\boldsymbol{R}_{r,0}}. \tag{A.18}$$

An R shock of size $\tilde{\varepsilon}_{r0}$ yields an interest rate response equal to $\boldsymbol{R}_{g,0}$. Next, we can evaluate $\mathbb{E}[\hat{c}_0]$ for an innovation to R of size $\tilde{\varepsilon}_{r0}$, instead of unity:

$$\begin{aligned} \mathbb{E}[\hat{c}_0 \times \hat{\tilde{\varepsilon}}_{r0}] &= \tilde{\gamma}_0 \boldsymbol{R}_{r,0}\tilde{\varepsilon}_{r0} \\ &= \tilde{\gamma}_0 \boldsymbol{R}_{g,0} \\ &= \Omega_0. \end{aligned} \tag{A.19}$$

The differential response of Y_{it} to an innovation to R of size $\tilde{\varepsilon}_{r0}$ equals the response that operates through the endogenous response of R to a G shock. In other words, the HEDGE term for $h = 0$.

For $h = 1$, it will be the case that, unless $\boldsymbol{R}_g$ and $\boldsymbol{R}_r$ are multiples of each other:

$$\boldsymbol{R}_{g,1} \neq \boldsymbol{R}_{r,1}\tilde{\varepsilon}_{r0}. \tag{A.20}$$

This means that the initial rescaling by $\tilde{\varepsilon}_{r0}$ is not enough to equalize the conditional paths of R at horizons $h > 0$. The next step is to feed in a contemporaneous innovation at date $h = 1$ that accounts for the remaining difference:

$$\tilde{\varepsilon}_{r1} = \boldsymbol{R}_{g,1} - \boldsymbol{R}_{r,1}\tilde{\varepsilon}_{r0} \tag{A.21}$$

$$= \boldsymbol{R}_{g,1} - \boldsymbol{R}_{r,1}\frac{\boldsymbol{R}_{g,0}}{\boldsymbol{R}_{r,0}}, \tag{A.22}$$

where $\tilde{\varepsilon}_{r1}$ is an innovation to R taking place at $t + 1$. This gives us two R shocks—$\{\tilde{\varepsilon}_{r0}, \tilde{\varepsilon}_{r1}\}$—that propagate to the cross-sectional outcome as follows

$$\mathbb{E}[\hat{c}_1 \times \hat{\tilde{\varepsilon}}_{r0} + \hat{c}_0 \times \hat{\tilde{\varepsilon}}_{r1}] = \left[\tilde{\gamma}_0 \boldsymbol{R}_{r,1} + \tilde{\gamma}_1 \boldsymbol{R}_{r,0}\right] \boldsymbol{R}_{g,0} + \tilde{\gamma}_0 \boldsymbol{R}_{r,0} \left[\boldsymbol{R}_{g,1} - \boldsymbol{R}_{r,1} \boldsymbol{R}_{g,0}\right]$$
$$= \tilde{\gamma}_0 \boldsymbol{R}_{r,1} \boldsymbol{R}_{g,0} + \tilde{\gamma}_1 \boldsymbol{R}_{g,0} + \tilde{\gamma}_0 \boldsymbol{R}_{g,1} - \tilde{\gamma}_0 \boldsymbol{R}_{r,1} \boldsymbol{R}_{g0} \tag{A.23}$$
$$= \tilde{\gamma}_1 \boldsymbol{R}_{g,0} + \tilde{\gamma}_0 \boldsymbol{R}_{g,1}$$
$$= \Omega_1.$$

In period $h = 1$, it is *as if* the economy was responding to a shock of size $\tilde{\varepsilon}_{r0}$ that took place a period ago plus to a shock of size $\tilde{\varepsilon}_{r1}$ taking place contemporaneously. The combination of these responses identifies Ω_1.

As before, for $h = 2$ we construct a contemporaneous shock at date $h = 2$ equivalent to the difference between $\boldsymbol{R}_{g,2}$ and the R path implied by the sequence of shocks fed in so far:

$$\tilde{\varepsilon}_{r2} = \boldsymbol{R}_{g,2} - \boldsymbol{R}_{r,2} \tilde{\varepsilon}_{r0} - \boldsymbol{R}_{r,1} \tilde{\varepsilon}_{r1}, \tag{A.24}$$

where $\boldsymbol{R}_{r,2}\tilde{\varepsilon}_{r0}$ is the response of R at $h = 2$ to the shock that took place at $t = 0$ and $\boldsymbol{R}_{r,1}\tilde{\varepsilon}_{r1}$ is the response of R at $h = 2$ to the shock that took place in the previous period. Note that because $R_{r,0} = 1$ I omit it from the left-hand side. Once we have $\tilde{\varepsilon}_{r2}$ we can construct an estimate for Ω_2 as:

$$\mathbb{E}[\hat{c}_2 \times \hat{\tilde{\varepsilon}}_{r0} + \hat{c}_1 \times \hat{\tilde{\varepsilon}}_{r1} + \hat{c}_0 \times \hat{\tilde{\varepsilon}}_{r2}] = \Omega_2. \tag{A.25}$$

The two formulas below generalize the computation of the sequence of shocks $\tilde{\varepsilon}_{rh}$ and the estimation of Ω_h for any $h \geq 0$.

$$\hat{v}_h = \sum_{k=0}^{h} \hat{a}_{h-k} \hat{\tilde{\varepsilon}}_{rk} \quad \forall\, h \in H, \tag{A.26}$$

$$\widehat{\Omega}_h = \sum_{k=0}^{h} \hat{c}_{h-k} \hat{\tilde{\varepsilon}}_{rk} \quad \forall\, h \in H. \tag{A.27}$$

A.1.3 Dynamic Setting without Anticipation Effects - Further Details

Control Function Approach in Dynamic Settings

This Appendix presents detailed derivations for the control function estimator.

The reason why the control function approach fails to control for the dynamic path of the GE variable in response to a shock of interest ε_{gt} is independent of whether the setting is a panel or a time-series regression. Given this, in this Appendix I provide details on the mechanics of the control function approach using an aggregate setting.

Throughout, I use the Frisch–Waugh–Lovell (FWL) theorem to characterize the expected value of the elasticities estimated under the control function approach. The FWL theorem implies:

$$\hat{b} = \frac{cov[y^{\perp z}, x^{\perp z}]}{v[x^{\perp z}]}, \tag{A.28}$$

where:

$$x^{\perp z} = x - \theta z, \qquad \theta = \frac{cov[x,z]}{v[z]}, \tag{A.29}$$

$$y^{\perp z} = y - \theta_y z, \qquad \theta_y = \frac{cov[y,z]}{v[z]}. \tag{A.30}$$

AR(1) Process for G and R

First, I consider the following DGP:

$$G_t = \rho_g G_{t-1} + u_{gt} \tag{A.31}$$

$$R_t = \rho_r R_{t-1} + \delta G_t + u_{rt} \tag{A.32}$$

$$Y_t = \beta G_t + \gamma R_t + u_{yt} \tag{A.33}$$

where I assume the researcher observes an innovation to G, ε_{gt}, that satisfies Assumption (1.3). The control function approach can be implemented with the following local projection:

$$Y_{t+h} = b_h G_t + c_h R_t + e_{t+h} \tag{A.34}$$

where G_t is instrumented with ε_{gt}. The FWL theorem implies the following expression for the 2SLS estimate $\hat{b}_h$:

$$\mathbb{E}[\hat{b}_h] = \frac{cov[Y_{t+h}^{\perp R_t}, \varepsilon_{gt}^{\perp R_t}]}{cov[G_t^{\perp R_t}, \varepsilon_{gt}^{\perp R_t}]}. \tag{A.35}$$

Then expanding and re-arranging the numerator yields:

$$\begin{aligned} cov[Y_{t+h}^{\perp R_t}, \varepsilon_{gt}^{\perp R_t}] &= \beta\rho_g^h\Big(cov[G_t, \varepsilon_{gt}] - \frac{cov[G_t, R_t]}{v[R_t]}cov[R_t, \varepsilon_{gt}]\Big) \\ &\qquad + \gamma\Big(cov[R_{t+h}, \varepsilon_{gt}] - \frac{cov[R_{t+h}, R_t]}{v[R_t]}cov[R_t, \varepsilon_{gt}]\Big) \\ &= \beta\rho_g^h\Big(v[\varepsilon_{gt}] - \delta\frac{cov[G_t, R_t]}{v[R_t]}v[\varepsilon_{gt}]\Big) \\ &\qquad + \gamma\Big(cov[R_{t+h}, \varepsilon_{gt}] - \frac{cov[R_{t+h}, R_t]}{v[R_t]}cov[R_t, \varepsilon_{gt}]\Big), \end{aligned} \tag{A.36}$$

where I used the fact that $cov[R_t, \varepsilon_{gt}] = \delta v[\varepsilon_{gt}]$. Similarly, the denominator can be expressed as:

$$cov[G_t^{\perp R_t}, \varepsilon_{gt}^{\perp R_t}] = v[\varepsilon_{gt}] - \delta\frac{cov[G_t, R_t]}{v[R_t]}v[\varepsilon_{gt}]. \tag{A.37}$$

Putting both pieces together, yields:

$$\mathbb{E}[\hat{b}_h] = \beta\rho_g^h + \frac{\gamma v[\varepsilon_{gt}^{\perp R_t}]}{cov[G_t^{\perp R_t}, \varepsilon_{gt}^{\perp R_t}]}\left(\boldsymbol{R}_{g,h} - \boldsymbol{R}_{r,h} \times \boldsymbol{R}_{\tilde{R},h}\right) \tag{A.38}$$

where:

$$\boldsymbol{R}_{\tilde{R},h} \equiv \frac{cov[R_{t+h}, R_t]}{v[R_t]}. \tag{A.39}$$

The term $\boldsymbol{R}_{\tilde{R},h}$ captures the component of the dynamic path of R that is predictable from the included control R_t. More generally, when additional controls are included, $\boldsymbol{R}_{\tilde{R},h}$ captures the projection of R_{t+h} onto the space spanned by those controls (e.g., current and lagged values of R and other aggregate variables).

This expression highlights that unless $\boldsymbol{R}_{g,h}$ can be represented as a linear combination of the paths of R spanned by included controls, the CFA fails to address the HEDGE bias.

Dynamic Setting - Monte Carlo Simulations

In the following, I detail the DGPs used to produce Table A.3 in the main body of the paper. For reference, I reproduce the general system of equations:

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \end{bmatrix},$$

$$Y_{it+h} = \sum_{s=0}^{S} \beta_s s_{ig} G_{t+h-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t+h-k} + \sum_{l=1}^{L} \psi_l Y_{i,t+h-l} + u_{ih} + u_{th} + u_{it+h}^{y} \quad \forall h \geq 0. \tag{A.40}$$

Each DGP is evaluated at three different values for the correlation between the exposure shifters s_{ig} and s_{ir}. I fix the standard deviation of all idiosyncratic shocks to unity. The sample size is set to $N = 300$, $T = 1000$, and the number of repetitions is set to $J = 1000$.

Table A.1. Monte Carlo DGPs — Dynamic Setting.

DGP #	**Description**	m_{gg}	m_{rr}	m_{gr}	m_{rg}	β_s	γ_k	ψ_l
1	VAR(1), no lagged Y_i	0.8	0.8	0.5	0.0	$\{1.0\}$	$\{0.5\}$	$\{0.0\}$
2	VAR(1), AR(1) in Y_i	0.8	0.8	0.5	0.0	$\{1.0\}$	$\{0.5\}$	$\{0.8\}$
3	VAR(2), AR(1) in Y_i	$\{0.7, 0.2\}$	$\{0.4, 0.5\}$	$\{0.5, 0.0\}$	$\{0.0, 0.0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.8\}$
4	VAR(1), lags in G and R in Y_i	0.7	0.4	0.5	0.0	$\{0.5, 0.2\}$	$\{0.5, 0.05\}$	$\{0.0\}$
5	VAR(1), unrestricted	0.8	0.8	0.5	-0.3	$\{1.0\}$	$\{0.5\}$	$\{0.0\}$

Note: Note: Each DGP is evaluated at $\phi \in \{0.1, 0.5, 1\}$. Parameters m_{xy} denote the effect of variable y on variable x in the VAR at different horizons. The β_s, γ_k, and ψ_l parameters correspond to the effects of G_{t-s}, R_{t-k}, and Y_{it-l} in the unit level outcome. Zero entries indicate either exclusion or no persistence.

Two-Way GE Feedback - Two Sources of Omitted Variable Bias

This appendix studies the cross-sectional identification of partial-equilibrium elasticities (PE) when there is two-way feedback (2W-GE) between aggregate variables. It shows that with generalized feedback between aggregate variables cross-sectional elasticities are generally contaminated by GE responses even when the covariance between exposure to the shock of interest ε_{gt} and exposure s_{ik} to some other aggregate variable K_t is zero.

The general setting is still characterized by (S2) - which I reproduce below for convenience:

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \end{bmatrix},$$

$$Y_{it} = \sum_{s=0}^{S} \beta_s s_{ig} G_{t-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t-k} + \sum_{l=1}^{L} \psi_l Y_{i,t-l} + u_i + u_t + u_{it}^{y} \tag{A.41}$$

$$s_{ir} = \phi s_{ig} + u_{is_r}.$$

Necessary Conditions to Difference-out GE using reduced form analysis

I study the necessary conditions for a reduced-form TWFE local projection to identify the partial equilibrium effect of a change in G_t on Y_{it+h}. Concretely, the regression under consideration is given by:

$$Y_{it+h} = b_h^{RF} s_{ig} \varepsilon_{gt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{A.42}$$

where ε_{gt} is an observed innovation to G in period t. Let b_h^{RF} be the coefficient of interest, where the superscript RF refers to reduced form. To build intuition, I start with the following simplified DGP for G and R:

$$G_t = \rho_g G_{t-1} + \alpha R_t + u_{gt}, \qquad R_t = \rho_r R_{t-1} + \delta G_t + u_{rt} \tag{A.43}$$

The VAR(1) representation can be written as

$$G_t = \frac{\rho_g}{1-\alpha\delta} G_{t-1} + \frac{\alpha\rho_r}{1-\alpha\delta} R_{t-1} + \frac{\alpha}{1-\alpha\delta} u_{rt} + \frac{1}{1-\alpha\delta} u_{gt} \tag{A.44}$$

$$R_t = \frac{\rho_r}{1-\alpha\delta} R_{t-1} + \frac{\delta\rho_g}{1-\alpha\delta} G_{t-1} + \frac{1}{1-\alpha\delta} u_{rt} + \frac{\delta}{1-\alpha\delta} u_{gt}. \tag{A.45}$$

Then, consider a simple DGP for Y_{it} given by:

$$Y_{it} = \beta s_{ig} G_t + \gamma s_{ir} R_t + u_{it}, \qquad s_{ir} = \phi s_{ig} + e_i^{s_r}. \tag{A.46}$$

It can be shown that the coefficient of interest converges to the following expression for $h = 0$ and $h = 1$, respectively:

$$\hat{b}_0^{RF} \to \frac{\beta}{1-\alpha\delta} + \frac{\phi\gamma\delta}{1-\alpha\delta}, \qquad \hat{b}_1^{RF} \to \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} + \frac{\phi\gamma\delta(\rho_g + \rho_r)}{(1-\alpha\delta)^2}. \tag{A.47}$$

Now, consider an economy in which exposure shifters to G and R are uncorrelated in the cross-section - that is, $\phi = 0$. This leaves us with the following.

$$\hat{b}_0 \to \frac{\beta}{1-\alpha\delta} \quad \neq \beta, \qquad \hat{b}_1 \to \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} \quad \neq \beta\rho_g. \tag{A.48}$$

Both reduced form coefficients are contaminated by GE effects even in absence of HEDGE. A necessary condition to *fully* difference-out GE using a reduced-form cross-sectional regression is that there should be no two-way feedback between the aggregate variable of interest and any variable that responds to it in general equilibrium. Otherwise, the reduced-form coefficients identify a rescaled version of the true partial effect even when $cov[s_{ig}, s_{ir}] = 0$.

Is 2SLS the solution?

Suppose that instead of using the reduced form, we chose to instrument G_t with ε_{gt}. The first stage coefficient is given by:

$$\theta^{FS} = \frac{1}{1-\alpha\delta}. \tag{A.49}$$

The 2SLS estimates converge to:

$$\hat{b}_0^{2S} \to \beta \quad = \beta, \qquad \hat{b}_1^{2S} \to \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)} \quad \neq \beta\rho_g. \tag{A.50}$$

For this DGP, using ε_{gt} as an instrument for G_t removes GE on impact but not at $h = 1$ (nor at any $h > 0$). What if instead we instrument G_{t+h} with ε_{gt}? In this case, the first stage coefficient is horizon-specific, and for $h = 1$, we get:

$$\theta_1^{FS} = \frac{\rho_g + \alpha\delta\rho_r}{(1-\alpha\delta)^2}. \tag{A.51}$$

The first stage coefficient for $h = 0$ is the same as in the previous case. The 2SLS estimates converge to:

$$\hat{b}_0^{2S,H} \to \beta, \qquad \hat{b}_1^{2S,H} \to \beta, \tag{A.52}$$

where I use the subscript H to denote that these estimates correspond to a 2SLS regression where G_{t+h} is being instrumented for, instead of G_t. In the context of this specific DGP, instrumenting G_{t+h} identifies the impact effect of a change in G and the local projection recovers the same coefficient at all horizons. Note that this does not identify the cumulative impulse response but does identify an object that is free of GE. Crucially, this worked because, for this DGP, G_{t+h}

captures, or blocks, all paths between ε_{gt} and Y_{it+h}. Consider departing from this scenario by adding an autoregressive component in the unit-level outcome:

$$Y_{it} = \beta s_{ig} G_t + \psi Y_{it-1} + u^y_{it}, \tag{A.53}$$

or, for lagged G to have an effect on the unit-level outcome:

$$Y_{it} = \beta s_{ig} G_t + \beta_1 s_{ig} G_{t-1} + u^y_{it}. \tag{A.54}$$

Now, G_{t+h} no longer blocks all channels through which a change in ε_{gt} affects Y_{it+h}. Therefore, instrumenting G_{t+h} with ε_{gt} will not be enough to difference-out GE effects. The 2SLS estimates for these two alternative DGPs at $h = 1$ are:

$$\hat{b}_1^{2S} = \beta \frac{(\rho_g + \delta\alpha\rho_r)}{1-\alpha\delta} + \beta\psi \ \neq \beta(\rho_g + \psi), \qquad \hat{b}_1^{2S} = \beta \frac{(\rho_g + \delta\alpha\rho_r)}{1-\alpha\delta} + \beta_1 \ \neq \beta\rho_g + \beta_1, \tag{A.55}$$

when instrumenting G_t, and

$$\hat{b}_1^{2S,H} = \beta + \beta\psi \frac{1-\alpha\delta}{(\rho_g + \delta\alpha\rho_r)} \ \neq \beta(\rho_g + \psi), \quad \hat{b}_1^{2S,H} = \beta + \beta_1 \frac{1-\alpha\delta}{(\rho_g + \delta\alpha\rho_r)} \ \neq \beta\rho_g + \beta_1, \tag{A.56}$$

when instrumenting G_{t+h}. The expressions to the right of the $\neq$ sign correspond to the PE elasticities for each DGP and horizon.

The Decomposition Framework also addresses this omitted variable bias

Using the reduced-form responses estimated from:

$$Y_{it+h} = b_h \times s_{ig}\varepsilon_{gt} + c_h \times s_{ig}\varepsilon_{rt} + _{ih} + \lambda_{th} + e_{it+h} \tag{A.57}$$

$$R_{t+h} = v_h \varepsilon_{gt} + a_h \varepsilon_{rt} + e_{rt+h}, \tag{A.58}$$

one can implement the decomposition framework as outlined in 1.3.2 and consistently estimate the partial equilibrium elasticity at all horizons, even when there is two-way feedback between aggregates in a system like $S2$. Next, I present results from a Monte Carlo simulation exercise and the algebraic details for the first two horizons in a simplified DGP.

HEDGE Test

In environments with 2W-GE feedback, the interpretation of the HEDGE test requires additional care. When G_t responds endogenously to R_t, variation in ε_{rt} induces changes in both aggregate variables. As a result, even in the absence of comovement between exposure shifters (i.e., $\phi = 0$), the interaction $s_{ig}\varepsilon_{rt}$ will be correlated with $s_{ig}G_t$.

In this case, a rejection of the null should not be interpreted as evidence of heterogeneous exposure to R per se (i.e., $\phi \neq 0$), but more generally as evidence of heterogeneous exposure to equilibrium responses of the economy. In particular, units that are more exposed to G will also be differentially exposed to the endogenous response of G to R. Therefore, the HEDGE test should be interpreted as a diagnostic for residual cross-sectional heterogeneity in exposure to aggregate responses, rather than a test exclusively of $\phi \neq 0$.

Monte Carlo Simulation - Results

For illustration, I ran a Monte Carlo exercise in which I simulate data from alternative DGPs that fit into the system given by (S2) to estimate b_h. Throughout, I set $\phi = 0$ to focus on the omitted variable bias that arises independently of HEDGE. The sample size is set to $N = 300$ and $T = 500$, the number of repetitions to $J = 1000$, and the standard deviation of all idiosyncratic shocks to unity. Table A.2 provides further details on each of the DGPs used in the Monte Carlo exercise. I consider four estimation strategies: (1) a TWFE reduced-form local projection, (2) the decomposition framework, (3) a TWFE local projection instrumenting G_t with ε_{gt}, and (4) a TWFE local projection instrumenting G_{t+h} with ε_{gt}. For each estimation strategy k, DGP z and horizon h, I compute the absolute difference between the true portable elasticity, β^P_{zh}, and the average estimate across all repetitions and then normalize it by β^P_{zh} to make it comparable across data-generating processes:

$$\theta^k_{zh} = \frac{\mid J^{-1} \sum_{j=1}^{J} \hat{b}^k_{jzh} - \beta^P_{zh} \mid}{\mid \beta^P_{zh} \mid}. \tag{A.59}$$

Consistency requires $\theta^k_{zh} \to 0$. Table A.3 shows the mean and standard deviation of θ^k_{zh} by horizon and estimation method. The decomposition approach outperforms all three estimation strategies at every horizon and yields unbiased estimates of the true portable elasticities, with only minor bias at longer horizons.

Table A.2. Two-Way GE - Parameterizations of DGPs in Monte Carlo Simulations.

DGP #	Description	$m_{gg} = m_{rr}$	m_{gr}	m_{rg}	β_s	γ_k	ψ_l
1	AR(1) in G and R, no lagged Y.	0.6	−0.3	0.5	{1.0}	{0.5}	{0}
2	Adds persistence in Y_{it} with one autoregressive lag.	0.6	0.1	0.5	{1.0}	{0.5}	{0.8}
3	Adds distributed lag response of Y to G via multiple β_s.	0.6	0.1	0.5	{1.0, 0.4, 0.2}	{0.5}	{0}
4	Combines lagged Y and lagged G.	0.6	0.1	0.5	{1.0, 0.4, 0.2}	{0.5}	{0.8}
5	Adds three autoregressive lags of Y.	0.6	0.1	0.5	{1.0}	{0.5}	{0.5, 0.3, −0.1}

Note: Parameters m_{xy} denote the effect of variable y on variable x in the VAR at different horizons. The β_s, γ_k, and ψ_l parameters correspond to the effects of G_{t_s}, R_{t-k}, and lags of Y_{it-l} in the unit level outcome. Zero entries indicate either exclusion or no persistence.

Table A.3. Absolute Mean Relative Bias by Estimation Method and Horizon.

Estimation Strategy	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
Decomposition Framework	0.002	0.005	0.009	0.015	0.019	0.027
	(0.001)	(0.002)	(0.006)	(0.010)	(0.010)	(0.016)
Reduced Form	0.002	0.078	0.211	0.391	0.587	0.788
	(0.002)	(0.098)	(0.273)	(0.469)	(0.636)	(0.744)
IV (t)	0.001	0.079	0.213	0.391	0.587	0.787
	(0.001)	(0.097)	(0.273)	(0.469)	(0.636)	(0.743)
IV (H)	0.001	0.624	1.556	3.494	5.159	8.680
	(0.001)	(0.026)	(0.191)	(1.662)	(1.123)	(2.353)

Note: Standard errors shown in parentheses. Values are reported as fractions of the true portable elasticity (1 = 100 %). The sample size is set to $N = 300$ and $T = 500$ and the number of repetitions to $J = 1000$.

Decomposition Framework with 2W-GE - Detailed Algebra for $h = 0, 1$

The DGP is the same as that used before, which assumes a VAR(1) process for aggregate variables and only contemporaneous G and R to directly affect the local outcome.

$$\begin{aligned} G_t &= \rho_g G_{t-1} + \alpha R_t + u_{gt}, \\ R_t &= \rho_r R_{t-1} + \delta G_t + u_{rt}, \\ Y_{it} &= \beta s_{ig} G_t + \gamma s_{ir} R_t + u_{it}^{y}, \\ s_{ir} &= \phi s_{ig} + u_i^{s_r}, \quad s_{ig} \perp u_i^{s_r}. \end{aligned} \tag{A.60}$$

Cross-sectional Estimates

The estimated coefficients of the cross-sectional step for $h = 0, 1$ converge to:

$$\mathbb{E}[\hat{b}_0] = \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta}, \qquad \mathbb{E}[\hat{b}_1] = \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}\delta(\rho_g + \rho_r)}{(1-\alpha\delta)^2}, \tag{A.61}$$

$$\mathbb{E}[\hat{c}_0] = \frac{\beta\alpha}{1-\alpha\delta} + \frac{\tilde{\gamma}}{1-\alpha\delta}, \qquad \mathbb{E}[\hat{c}_1] = \frac{\beta\alpha(\rho_g + \rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}(\rho_r + \delta\alpha\rho_g)}{(1-\alpha\delta)^2}. \tag{A.62}$$

Time Series Estimates

The estimated coefficients from the time series step for $h = 0, 1$ converge to:

$$\mathbb{E}[\hat{a}_0] = \frac{1}{1-\alpha\delta}, \qquad \mathbb{E}[\hat{a}_1] = \frac{\rho_r + \delta\alpha\rho_g}{(1-\alpha\delta)^2}, \tag{A.63}$$

$$\mathbb{E}[\hat{v}_0] = \frac{\delta}{1-\alpha\delta}, \qquad \mathbb{E}[\hat{v}_1] = \frac{\delta(\rho_r + \rho_g)}{(1-\alpha\delta)^2}. \tag{A.64}$$

Decomposition

The counterfactual shock for $h = 0$ satisfies the following:

$$v_0 = a_0 \tilde{\varepsilon}_{r0} \rightarrow \frac{\delta}{1-\alpha\delta} = \frac{1}{1-\alpha\delta}\tilde{\varepsilon}_{r0} \tag{A.65}$$

$$\delta = \tilde{\varepsilon}_{r0}. \tag{A.66}$$

Next, we can evaluate the differential response to R at $\tilde{\varepsilon}_{r0}$ and subtract it from the differential response to G:

$$\mathbb{E}[\hat{b}_0 - \hat{c}_0 \times \hat{\tilde{\varepsilon}}_{r0}] = \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta} - \left[\frac{\beta\alpha}{1-\alpha\delta} + \frac{\tilde{\gamma}}{1-\alpha\delta}\right] \times \delta \tag{A.67}$$

$$= \frac{\beta(1-\alpha\delta)}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta} - \frac{\tilde{\gamma}\delta}{1-\alpha\delta} \tag{A.68}$$

$$= \beta. \tag{A.69}$$

In period $h = 1$, this shock $\tilde{\varepsilon}_{r0} = \delta$ implies the following realization for R_{t+1}:

$$\boldsymbol{R}_{r,1} = \frac{\rho_r + \delta\alpha\rho_g}{(1-\alpha\delta)^2}\delta \quad \neq \quad \boldsymbol{R}_{g,1} = \frac{\delta(\rho_r + \rho_g)}{(1-\alpha\delta)^2}. \tag{A.70}$$

We want to find a second date $h = 1$ shock $\tilde{\varepsilon}_{r1}$ such that:

$$\begin{aligned}
\mathbb{E}[\hat{v}1] &= \mathbb{E}[\hat{a}_1 \times \hat{\tilde{\varepsilon}}_{r0} + \hat{a}_0] \times \hat{\tilde{\varepsilon}}_{r1}] \\
\frac{\delta(\rho_r + \rho_g)}{(1-\alpha\delta)^2} &= \frac{\rho_r + \delta\alpha\rho_g}{(1-\alpha\delta)^2}\delta + \frac{1}{1-\alpha\delta} \times \mathbb{E}[\hat{\tilde{\varepsilon}}_{r1}] \\
&\rightarrow \mathbb{E}[\hat{\tilde{\varepsilon}}_{r1}] = \delta\rho_g.
\end{aligned} \tag{A.71}$$

Next, we evaluate the differential response of Y_i to these two R shocks, $\{\tilde{\varepsilon}_{r0}, \tilde{\varepsilon}_{r1}\}$:

$$\mathbb{E}[\hat{c}_1 \times \hat{\tilde{\varepsilon}}_{r0} + \hat{c}_0 \times \hat{\tilde{\varepsilon}}_{r1}], \tag{A.72}$$

and subtract it from the differential response of the cross-sectional outcome to a G shock in period $t + 1$:

$$\begin{aligned}
\mathbb{E}[\hat{b}_1 - \hat{c}_1 \times \hat{\tilde{\varepsilon}}_{r0} - \hat{c}_0 \times \hat{\tilde{\varepsilon}}_{r1}] &= \left[\frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}\delta(\rho_g + \rho_r)}{(1-\alpha\delta)^2}\right] \\
&\quad - \left[\frac{\beta\alpha(\rho_g + \rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}(\rho_r + \delta\alpha\rho_g)}{(1-\alpha\delta)^2}\right]\delta - \frac{\beta\alpha + \tilde{\gamma}}{1-\alpha\delta}\rho_g\delta \\
&= \beta\rho_g = \beta_1^P.
\end{aligned} \tag{A.73}$$

This process can be extended to further horizons.

Robustness to $cov[\varepsilon_{gt}, A_t] \neq 0$

This appendix lifts the assumption that the researcher observes a purely structural innovation to G_t. I consider settings in which ε_{gt}, the observed shock of interest, is correlated with the realization of a third aggregate variable A_t. This nests cases in which the shock is not a source of exogenous variation at the aggregate level but satisfies exogeneity in a cross-sectional setting, conditional on the exposure shifters and time fixed effects. For illustration purposes, consider the following minimal setting.

$$G_t = \rho_g G_{t-1} + u_{gt}, \tag{A.74}$$
$$R_t = \rho_r R_{t-1} + \delta G_t + u_{rt}, \tag{A.75}$$
$$A_t = \kappa_a G_t + u_{at}, \tag{A.76}$$
$$Y_{it} = \beta s_{ig} G_t + \gamma s_{ir} R_t + \chi s_{ia} A_t + u_i + u_t + u_{it}^{y}, \tag{A.77}$$
$$s_{ir} = \phi s_{ig} + u_i^{s_r}, \tag{A.78}$$
$$cov[s_{ig}, s_{ia}] = 0. \tag{A.79}$$

Here, A_t is a third aggregate variable that contemporaneously responds to changes in G_t and to its own structural innovation. In addition, units are differentially exposed to A according to the exposure shifter s_{ia} and the parameter χ. In a cross-sectional TWFE regression, the shock ε_{gt} satisfies exogeneity with respect to A_t if $cov[s_{ig}, s_{ia}) = 0$. That is, the cross section identifies a causal effect of ε_{gt} if there is no HEDGE in relation to A_t.

Next, I show that the decomposition framework consistently identifies the portable elasticity in settings that use variation correlated to a third aggregate variable. I consider DGPs that fit into the following system of equations:

$$\begin{bmatrix} G_t \\ R_t \\ A_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \\ A_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \\ u_{at} \end{bmatrix},$$

$$\begin{aligned} Y_{it} = & \sum_{s=0}^{S} \beta_s s_{ig} G_{t-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t-k} + \sum_{w=0}^{W} \chi_w s_{ia} A_{t-w}, \\ & + \sum_{l=1}^{L} \psi_l Y_{i,t-l} + u_i + u_t + u_{it}^{y}, \\ s_{ir} = & \ \phi s_{ig} + u_i^{s_r}, \qquad s_{ia} = \xi s_{ig} + u_i^{s_a}, \qquad s_{ig} \sim \mathcal{N}(1, u_{s_g}). \end{aligned} \tag{S3}$$

For simulations, I study DGPs that (i) satisfy the identification assumption of the researcher - namely $cov[s_{ig}, s_{ia}] = 0$ ($\xi = 0$); and (ii) fail to satisfy it because they display HEDGE in A_t, on top of R_t. In both cases, the decomposition identifies its object of interest. That is, a sequence of elasticities that are clean of the effects operating through changes in R. The difference is that in the first case, the estimated elasticities are also independent of changes in A_t, whereas in the second case they are a combination of the true portable elasticity and the HEDGE channel associated with A_t. Table A.4 provides details on each of the DGPs used in the Monte Carlo exercise. In all cases, the sample size is set to $N = 100$ and $T = 500$, the number of repetitions to $J = 500$, and the standard deviation of all idiosyncratic shocks to unity. For each estimation strategy k, DGP z and horizon h, I compute the absolute difference between the true portable elasticity, β^P_{zh} and the average estimate across all repetitions and then normalize it by β^P_{zh} to make it comparable across data-generating processes:

$$\theta^k_{zh} = \frac{\mid J^{-1} \sum_{j=1}^{J} \hat{b}^k_{jzh} - \beta^P_{zh} \mid}{\mid \beta^P_{zh} \mid}. \tag{A.80}$$

Consistency requires $\theta^k_{zh} \to 0$. Table A.5 shows the mean and standard deviation of θ^k_{zh} by horizon and estimation method. The decomposition approach consistently recovers the true portable effects as opposed to the control function approach.

Table A.4. Three-Way GE — Parameterizations of DGPs in Monte Carlo Simulations.

DGP - Description	m_{gg}	m_{gr}	m_{ga}	m_{rr}	m_{rg}	m_{ra}	m_{aa}	m_{ag}	m_{ar}	β_s	γ_k	χ_w	ψ_l	ξ
A is i.i.d..	0.9	$\{0,0\}$	$\{0,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.0	$\{0.2\}$	$\{0,0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds persistence in A_t.	0.9	$\{0,0\}$	$\{0,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.4	$\{0.2\}$	$\{0,0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds $R \to A$ channel via m_{ar}.	0.9	$\{0,0\}$	$\{0,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.4	$\{0.2\}$	$\{0.4,0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds $A \to G$ channel via m_{ga}.	0.8	$\{0,0\}$	$\{0.1,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.4	$\{0.2\}$	$\{0.2,0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
3W-GE.	0.9	$\{-0.3,0\}$	$\{0.1,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.4	$\{0.2\}$	$\{0.2,0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds lags of G, R, and A in Y_{it}.	0.9	$\{-0.3,0\}$	$\{0.1,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.4	$\{0.2\}$	$\{0.2,0\}$	$\{1.0,0.2\}$	$\{0.5,0.2\}$	$\{0.2,0.3\}$	0.8	0
Adds HEDGE-A via $\xi > 0$.	0.9	$\{-0.3,0\}$	$\{0.1,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.4	$\{0.2\}$	$\{0.2,0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0.5
Combines lagged inputs with HEDGE-A.	0.9	$\{-0.3,0\}$	$\{0.1,0\}$	0.7	$\{0.5,0\}$	$\{0.3,0\}$	0.4	$\{0.2\}$	$\{0.2,0\}$	$\{1.0,0.2\}$	$\{0.5,0.2\}$	$\{0.2,0.3\}$	0.8	0.5

Note: Each DGP is evaluated at $\phi \in \{0.1, 0.5, 1\}$. Parameters m_{xy} denote the effect of variable y on variable x in the VAR at different horizons. The β_s, γ_k, χ_w and ψ_l parameters correspond to the effects of G_{t-s}, R_{t-k}, A_{t-j} and Y_{it-l} in the unit level outcome. Zero entries indicate either exclusion or no persistence.

Table A.5. DGPs with $cov[A_t, \varepsilon_{gt}] \neq 0$ - Absolute Mean Relative Bias by Estimation Method.

Estimation Method	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$
Decomposition	0.004	0.005	0.007	0.009	0.010	0.011	0.012
	(0.003)	(0.003)	(0.005)	(0.006)	(0.007)	(0.008)	(0.010)
Control Function	1.266	0.560	0.331	0.236	0.246	0.343	0.453
	(0.414)	(0.167)	(0.073)	(0.129)	(0.189)	(0.202)	(0.210)

Note: Standard errors in parenthesis. Values are reported as fractions of the true portable elasticity (1 = 100 %). Based on $J = 500$ repetitions with $N = 100$ and $T = 500$.

A.1.4 Comparison with Other Estimation Approaches

This appendix discusses how the decomposition approach outlined in Section 1.3.2 compares to three alternative approaches: (i) Interactive fixed effects (Bai, 2009), (ii) control function approach assuming s_{ir} is observed, (iii) 2SLS control function approach.

Interactive Fixed Effects

The cross-sectional outcome Y_{it} evolves according to

$$Y_{it} = \beta \times s_{ig} G_t + \gamma s_{ir} R_t + u_i + u_t + u^y_{it} \tag{A.81}$$

The Interactive Fixed Effects (IFE) estimator (Bai, 2009) augments the TWFE regression with L latent factors to absorb unobserved heterogeneity:

$$Y_{it+h} = \beta_h^{\text{IFE}} \times s_{ig}\varepsilon_{gt} + \sum_{\ell=1}^{r} \lambda_i^\ell f_t^\ell + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{A.82}$$

where $\{\lambda_i^\ell, f_t^\ell\}$ are estimated jointly with β_h^{IFE} via alternating least squares (ALS). The rationale is that if the residual error contains a factor structure orthogonal to $s_{ig}\varepsilon_{gt}$, the ALS procedure recovers both the factor and the causal coefficient without bias. After removing two-way fixed effects, the residual contains the GE term:

$$\eta_{it+h} = \gamma s_{ir} R_{t+h} + u^y_{it+h}. \tag{A.83}$$

From (S2), the residual satisfies:

$$\eta_{it+h} = \gamma s_{ir} \underbrace{\left[\psi_h^R(\varepsilon_{gt}) + \psi_h^R(\varepsilon_{rt})\right]}_{R_{t+h}} + u^y_{it+h} = \underbrace{\gamma s_{ir}}_{\lambda_i} \underbrace{\left[\psi_h^R(\varepsilon_{gt}) + \psi_h^R(\varepsilon_{rt})\right]}_{f_t} + u^y_{it+h} \tag{A.84}$$

where $\psi_h^R(\varepsilon_{kt})$ denotes the component of R_t driven by contemporaneous and past ε_k shocks for $k = g, r$. The residual admits a rank-1 factor structure, as all components of R_{t+h} load on the same cross-sectional vector s_{ir}. However, this factor is a function of both ε_{gt} and ε_{rt}, and is therefore correlated with the regressor $s_{ig}\varepsilon_{gt}$. As a consequence, the ALS algorithm cannot separate the variation in Y_{it+h} generated by $s_{ig}\varepsilon_{gt}$ from that generated by $\gamma s_{ir}\psi_h^R(\varepsilon_{kt})$, because the two share the same cross-sectional loading direction and their time-series components are correlated.

With a single factor, the IFE model is correctly specified in terms of rank. However, the factor $f_t = \psi_h^R(\varepsilon_{gt}) + \psi_h^R(\varepsilon_{rt})$ is itself a function of ε_{gt} and therefore correlated with the regressor $s_{ig}\varepsilon_{gt}$ in both time-series and cross-sectional dimensions. As a result, the orthogonality condition required for IFE to recover the causal coefficient fails. In the ALS procedure, the regressor and the factor compete to explain the same variation in Y_{it+h}, leading the estimator to attribute part of the treatment effect to the factor and producing an overcorrection relative to TWFE.

Figure A.2 presents results from a Monte Carlo simulation to illustrate the previous discussion.

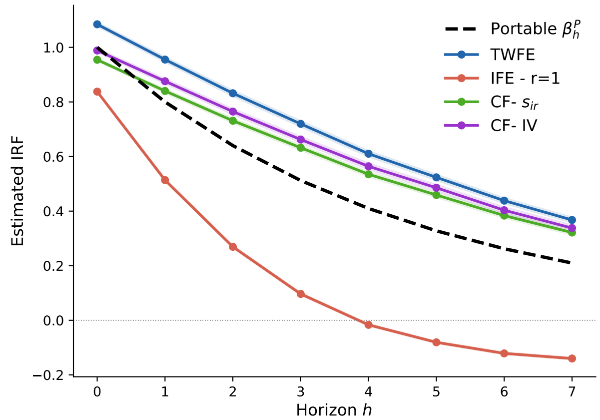


Figure A.2. Monte Carlo Simulations - Alternative Estimation Approaches.

Note: Simulation results based on 1000 repetitions with $N = 100$ and $T = 300$. 90% confidence bands. The DGP is generated using the following parameter values: $\rho_g = \rho_r = .8, \delta = .5, \beta = 1, \gamma = .5, \phi = .5$. The variance of all innovations is set to unity.

Controlling for $s_{ir}R_t$

Suppose that we had knowledge of the exposure shifters s_{ir} and extended the baseline specification with an interaction between these and R_t:

$$Y_{it+h} = \beta_h^{CF-s_{ir}} \times s_{ig}\varepsilon_{gt} + \gamma_h \times s_{ir}R_t + \lambda_{ih} + \lambda_{th} + e_{it+h} \tag{A.85}$$

This strategy fails to recover the true portable elasticity under the same conditions as the control function approach. The interaction $s_{ir}R_t$ exploits the same time-series variation as the control function and therefore cannot separately identify the component of R_t driven by ε_{gt}. As a result, it does not control for the endogenous response of R_t to ε_{gt} unless the corresponding time paths are linearly dependent.

This extension highlights that the difficulty in addressing the bias associated with HEDGE stems from the dynamic response of macroeconomic aggregates. Figure A.2 presents results from a Monte Carlo simulation to illustrate this point.

Controlling for R_t instrumented with ε_{rt}

This strategy differs from the control function approach by instrumenting R_t rather than using its realized value. By a similar argument, it also fails. While the control function uses the variation in R_t orthogonal to ε_{gt}, this approach isolates the component of R_t driven by ε_{rt}. Unless the path of R_t in response to ε_{rt} spans its response to ε_{gt}, this variation is insufficient to account for the endogenous response of R_t to ε_{gt}. Figure A.2 presents results from a Monte Carlo simulation to illustrate this point.

A.1.5 Forward-Looking Settings - Further Details

Figure A.3 shows the response of R to ε_{gt}, the fiscal shock, and to each of the two scenarios for the persistence of ε_{it}, the monetary shock.

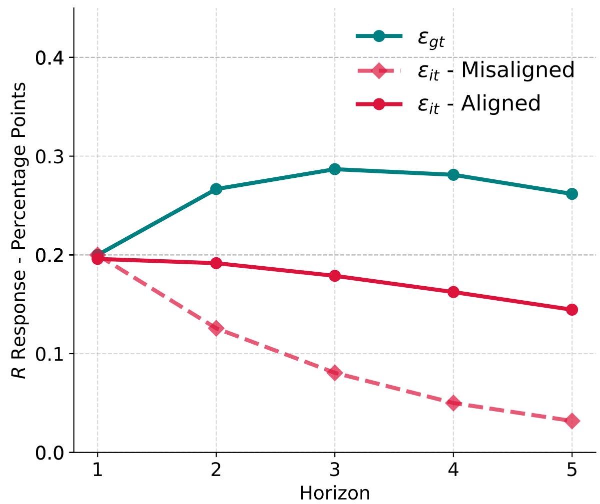


Figure A.3. Path of R Conditional on Aggregate Shocks.

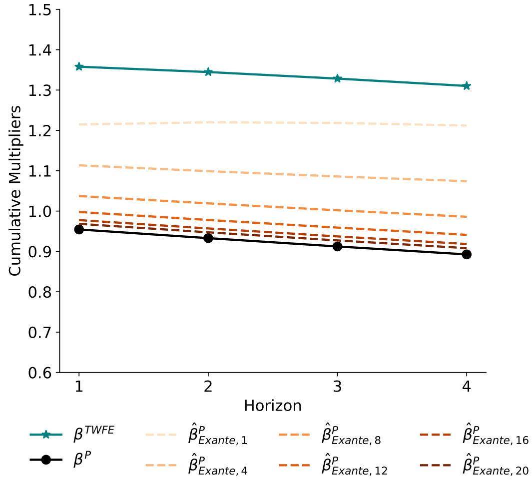
Figure A.4 presents additional simulation results. The upper row is an extension of Figure 1.3 in the main text that considers additional news shocks. It can be seen that, both in the *Misaligned* and *Aligned*, incorporating further news shocks improves the performance of the decomposition until it finally recovers the true portable elasticities. In the *Aligned* case, using news beyond $\tilde{F} = 4$ yields negligible gains.

The bottom row considers an economy with weaker forward-looking dynamics by introducing a discount term in the Euler equation[1] of households (McKay et al., 2017). Discounting weakens the importance of expected output gaps for current consumption responses and, overall, moves the economy closer to the Markov setting. I set the discounting parameter to $\alpha = .95$ (relative to a baseline without discounting, $\alpha = 1$). As expected, decreasing the strength of forward-looking dynamics improves the performance of a decomposition that uses a limited number of news shocks.

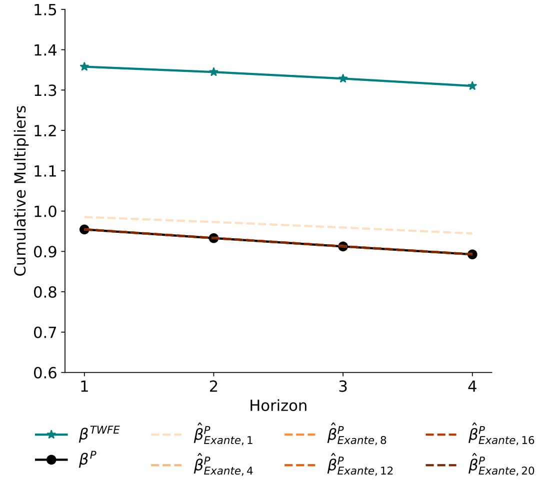
[1]The Euler equation is given by:

$$C_t^{-\frac{1}{\sigma}} = \vartheta\beta\mathbb{E}_t\left[C_{t+1}{}^{-\frac{\alpha}{\sigma}}(1+r_{t+1})\right], \tag{A.86}$$

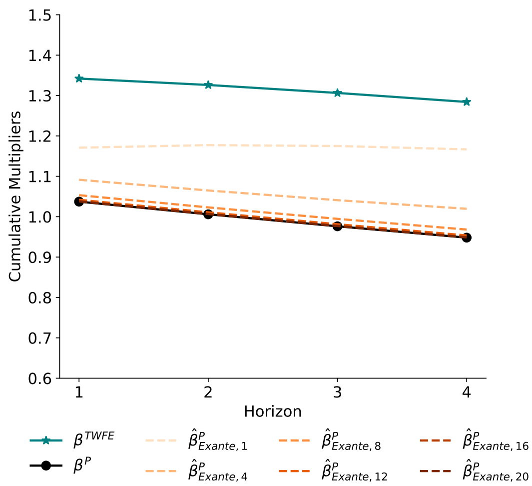
where α is the discounting parameter and ϑ is a constant calibrated to match steady-state consumption.



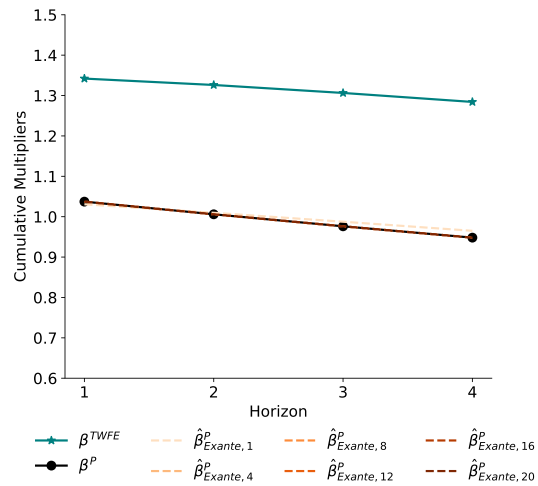
(A) *Misaligned R* Path - No Discounting.



(B) *Aligned R* Path - No Discounting.



(C) *Misaligned R* Path - Discounting.



(D) *Aligned R* Path - Discounting.

Figure A.4. Forward-looking Settings - Additional Simulation Results.

Note: *Misaligned* refers to the case where the trajectory of R differs across shocks, while *Aligned* corresponds to the case where these trajectories are relatively similar. The top row shows the estimation results using different numbers of news shocks in an economy without discounting (i.e., standard Euler equation), while the bottom row corresponds to an economy with a discounted Euler equation (McKay et al., 2017). The discounting parameter is set to $\alpha = .95$.

A.1.6 Asymptotic Properties

This section establishes the consistency of the estimators for β^P_h. The argument proceeds in three steps. First, I establish consistency of the panel local projection estimators under large-N, fixed-T asymptotics. Second, I establish consistency of the time-series local projection estimators under large-T asymptotics. Third, I show that $\hat{\beta}^P_h$ is a continuous mapping of these estimators and therefore consistent by the Continuous Mapping Theorem under suitable non-degeneracy conditions.

Assumptions for Panel Local Projections

Throughout, I consider asymptotics as $N \to \infty$ for T fixed. Aggregate shocks and time fixed effects are treated as fixed in the cross-sectional limit.

Assumption A.1.

$$\mathbb{E}_i \left[s_{ig} e_{it+h} \mid \boldsymbol{X} \right] = 0,$$

where $\boldsymbol{X} = \varepsilon_{gt}, \varepsilon_{rt}, \lambda_{ih}, \lambda_{th}$. This is the key identifying condition for a panel TWFE local projection (that relies on cross-sectional variation for identification). It holds if, conditional on the realization of aggregate shocks, the fixed effects and included controls, high and low exposure units do not experience systematically different unobserved shocks.

Assumption A.2.

$$\sup_{i,t} \mathbb{E}\left[e_{it+h}^2 \mid \boldsymbol{X} \right] \leq \sigma^2 < \infty, \quad \sup_i s_{ig}^2 < \infty$$

Assumption A.3. *For each t and h,*

$$\frac{1}{N^2} \sum_{i \neq j} s_{ig} s_{jg} cov[e_{it+h}, e_{jt+h} \mid \boldsymbol{X}] \to 0 \quad as\ N \to \infty.$$

This condition allows for common factors with heterogeneous loadings and/or cluster components. It requires that exposure shares are sufficiently dispersed within and across clusters, and asymptotically orthogonal to heterogeneous loadings (i.e. no HEDGE). This assumption does not preclude arbitrary serial correlation within units i (i.e. $cov[e_{it}, e_{is} \mid \boldsymbol{X}]$ is unrestricted for $s \neq t$).

Assumption A.4.

$$\frac{1}{N} \sum_{i=1}^{N} \tilde{s}_{ig}^2 \to \mu_{s^2} > 0 \quad as\ N \to \infty,$$

where $\tilde{s}_{ig}$ denotes the within-transformed exposure. This ensures that cross-sectional variation in exposures does not degenerate and is an identification condition that allows b_h and c_{hf} to be separately identified from λ_{th}.

Assumption A.5. *The aggregate shocks satisfy:*

$$\sum_{t=1}^{T} \tilde{\varepsilon}_{gt} \tilde{\varepsilon}_{rt+f} = 0 \quad \forall f = 0, \ldots, F, \qquad and \qquad \sum_{t=1}^{T} \tilde{\varepsilon}_{rt+f} \tilde{\varepsilon}_{rt+f'} = 0 \quad \forall f \neq f'.$$

This holds exactly when all shock series are mutually orthogonal in the time-series dimension, which can be imposed without loss of generality by orthogonalizing the shocks prior to estimation. Assumption A.5 is made to simplify the proof exposition because it allows the multivariate regression to be analyzed via scalar OLS formulas.

After within transformation:

$$\tilde{y}_{it+h} = b_h \tilde{s}_{ig} \tilde{\varepsilon}_{gt} + \sum_{f=0}^{F} c_{h,f} \tilde{s}_{ig} \tilde{\varepsilon}_{rt+f} + \tilde{e}_{it+h}, \tag{A.87}$$

where $\tilde{z}$ denotes the within transformation of variable z.

Under Assumption A.5, the regressors $\tilde{s}_{ig}\tilde{\varepsilon}_{gt}$ and $\{\tilde{s}_{ig}\tilde{\varepsilon}_{rt+f}\}_{f=0}^{F}$ are mutually orthogonal in the sample, so the TWFE estimators reduce to:[2]

$$\hat{b}_h = b_h + \frac{\sum_{t,i} \tilde{s}_{ig} \tilde{\varepsilon}_{gt} \tilde{e}_{it+h}}{\sum_{t,i} (\tilde{s}_{ig} \tilde{\varepsilon}_{gt})^2}, \tag{A.88}$$

$$\hat{c}_{hf} = c_{hf} + \frac{\sum_{t,i} \tilde{s}_{ig} \tilde{\varepsilon}_{rt+f} \tilde{e}_{it+h}}{\sum_{t,i} (\tilde{s}_{ig} \tilde{\varepsilon}_{rt+f})^2}. \tag{A.89}$$

For fixed T and $N \to \infty$, it suffices to show that[3]

$$\frac{1}{NT} \sum_{t,i} (\tilde{s}_{ig} \tilde{\varepsilon}_{kt})^2 \to C > 0 \;\; \forall\, k = g, r, \tag{A.91}$$

$$A_{Nth} = \frac{1}{N} \sum_{i} \tilde{s}_{ig} \tilde{e}_{it+h} \xrightarrow{p} 0 \;\text{ for each } t \text{ and } h. \tag{A.92}$$

For the first condition, since T is fixed and $\tilde{\varepsilon}_{kt}$ is constant across i,

$$\frac{1}{NT} \sum_{t,i} (\tilde{s}_{ig} \tilde{\varepsilon}_{kt})^2 = \underbrace{\frac{1}{T} \sum_{t} \tilde{\varepsilon}_{kt}^2}_{>0 \text{ a.s.}} \cdot \frac{1}{N} \sum_{i} \tilde{s}_{ig}^2 \;\longrightarrow\; \frac{1}{T} \sum_{t} \tilde{\varepsilon}_{kt}^2 \cdot \mu_{s^2} \;>\; 0, \tag{A.93}$$

where the first factor is strictly positive almost surely provided the shocks are non-degenerate, and the second factor converges to a strictly positive constant by Assumption A.4.

[2]The general multivariate case without Assumption A.5 requires positive definiteness of the matrix $Z'Z/NT$. Because all regressors share the same cross-sectional loading $\tilde{s}_{ig}$, this matrix converges to $\mu_{s^2} \cdot \Sigma_{\varepsilon}$, where Σ_{ε} is the variance-covariance matrix of the shock series. Positive definiteness therefore holds if and only if Assumption A.4 holds and the shocks are linearly independent over the T observed periods. The latter is assumed in the time-series step of the proof.

[3]The condition for the numerator follows from the fact that the numerator can be factored as:

$$\frac{1}{NT} \sum_{t,i} \tilde{s}_{ig} \tilde{\varepsilon}_{kt} \tilde{e}_{it+h} = \frac{1}{T} \sum_{t} \tilde{\varepsilon}_{kt} \underbrace{\left(\frac{1}{N} \sum_{i} \tilde{s}_{ig} \tilde{e}_{it+h} \right)}_{= A_{Nth}}, \tag{A.90}$$

where the factoring uses the fact that $\tilde{\varepsilon}_{kt}$ is constant across i and fixed conditional on $\boldsymbol{X}$. Since T is fixed, it suffices to show $A_{Nth} \xrightarrow{p} 0$ for each t and h; the full numerator is then a finite weighted sum of $o_p(1)$ terms and hence $o_p(1)$.

For the second condition, it suffices to show that:

$$\mathbb{E}[A_{Nth} \mid \boldsymbol{X}] = 0 \quad \text{and} \quad \mathbb{V}[A_{Nth} \mid \boldsymbol{X}] \to 0, \tag{A.94}$$

The first condition follows from Assumptions A.1 and A.2. Since the within transformation is a bounded linear operation in T, it preserves conditional mean independence.[4] The conditional variance of A_{Nth} is given by:

$$\mathbb{V}[A_{Nth} \mid \boldsymbol{X}] = \frac{1}{N^2}\left(\sum_i \tilde{s}_{ig}^2 \mathbb{V}[\tilde{e}_{it+h} \mid \boldsymbol{X}] + \sum_{i \neq j} \tilde{s}_{ig}\tilde{s}_{jg} cov[\tilde{e}_{it+h}, \tilde{e}_{jt+h} \mid \boldsymbol{X}]\right) \tag{A.95}$$

Weak cross-sectional dependence, as stated in A.3, implies that the second term in brackets goes to zero asymptotically. Assumptions A.4 and A.2 imply[5] that the first term goes to zero too, hence:

$$\mathbb{V}[A_{Nth} \mid \boldsymbol{X}] \leq \frac{C\sigma^2}{N}\frac{1}{N}\sum_i \tilde{s}_{ig}^2 = O\left(\frac{1}{N}\right) \to 0 \; as \; N \to \infty \tag{A.96}$$

Then, by Chebyshev's inequality:

$$P(|A_{Nth} - \mathbb{E}[A_{Nth} \mid \boldsymbol{X}]| > \varepsilon \mid \boldsymbol{X}) \leq \frac{\mathbb{V}[A_{Nth} \mid \boldsymbol{X}]}{\varepsilon^2} \to 0. \tag{A.97}$$

Since $\mathbb{E}[A_{Nth} \mid \boldsymbol{X}] = 0$, we have $|A_{Nth} - \mathbb{E}[A_{Nth} \mid \boldsymbol{X}]| = |A_{Nth}|$, and by the law of total probability:

$$P(|A_{Nth}| > \varepsilon) = \mathbb{E}[P(|A_{Nth}| > \varepsilon \mid \boldsymbol{X})] \to 0. \tag{A.98}$$

Therefore:

$$\hat{b}_h - b_h \xrightarrow{p} 0, \qquad \hat{c}_{hf} - c_{hf} \xrightarrow{p} 0 \quad \forall\, h \text{ and } \forall\, f. \tag{A.99}$$

Assumptions for Time-series Local projections

Assumption A.6. $\{\{\varepsilon_{rt+f}\}_{f=0}^{F}, \varepsilon_{gt}, e_{t+h}\}$ *is covariance-stationary and ergodic with:*

$$\mathbb{E}[\varepsilon_{rt+f}] = \mathbb{E}[\varepsilon_{gt}] = \mathbb{E}[e_{t+h}] = 0$$

Assumption A.7.

$$\mathbb{E}[\varepsilon_{rt+f}^2] = \sigma_{r,f}^2 < \infty, \mathbb{E}[\varepsilon_{gt}{}^2] = \sigma_g^2 < \infty, \quad \mathbb{E}[e_{t+h}^2] = \sigma_e^2 < \infty$$

[4] The within transformation writes $\tilde{e}_{it+h} = e_{it+h} - \bar{e}_{i\cdot} - \bar{e}_{\cdot t+h} + \bar{e}_{\cdot\cdot}$, where each cross-sectional average $\bar{e}_{\cdot t+h} = \frac{1}{N}\sum_j e_{jt+h}$ satisfies $\mathbb{E}[\tilde{s}_{ig}\bar{e}_{\cdot t+h} \mid \boldsymbol{X}] = 0$ by Assumption A.1 applied uniformly across j.

[5] The within transformation preserves the boundedness of second moments.

Assumption A.8.

$$\mathbb{E}[\varepsilon_{rt+f}^2] = \sigma_{r,f}^2 > 0, \quad \mathbb{E}[\varepsilon_{gt}^2] = \sigma_g^2 > 0$$

This condition guarantees that the shocks have variation over time.

Assumption A.9.

$$\mathbb{E}[\varepsilon_{rt+f}e_{t+h}] = 0, \quad \mathbb{E}[\varepsilon_{gt}e_{t+h}] = 0$$

The OLS estimator $(\hat{\nu}_h, \{\hat{a}_{h,f}\}_{f=0}^F)$ satisfies:

$$\begin{pmatrix} \hat{\nu}_h - \nu_h \\ \{\hat{a}_{hf} - a_{hf}\}_{f=0}^F \end{pmatrix} = \left(\frac{1}{T}\sum_t \mathbf{z}_t\mathbf{z}_t^\top\right)^{-1} \frac{1}{T}\sum_t \mathbf{z}_t\, e_{t+h}$$

where $\mathbf{z}_t = (\varepsilon_{gt}, \{\varepsilon_{rt+f}\}_{f=0}^F)^\top$. By Assumption A.6 and A.8, as $T \to \infty$, $\frac{1}{T}\sum_t \mathbf{z}_t\mathbf{z}_t^\top \xrightarrow{p} \Sigma_{zz}$, which is positive definite under Assumption A.8. By Assumptions A.9 and A.7 and the ergodic LLN, $\frac{1}{T}\sum_t \mathbf{z}_t e_{t+h} \xrightarrow{p} \mathbb{E}[\mathbf{z}_t e_{t+h}] = \mathbf{0}$. Slutsky's theorem gives $(\hat{\nu}_h - \nu_h, \{\hat{a}_{hf} - a_{hf}\}_{f=0}^F)^\top \xrightarrow{p} \mathbf{0}$.

Consistency of $\hat{\beta}_h^P$

For the Markov setting, the portable elasticity at horizon h is defined as:

$$\beta_h^P = g^M\Big(b_h, \{c_s, \nu_s, a_s\}_{s=0}^H\Big) \tag{A.100}$$

where g^M is a continuous mapping on the domain $\Theta_h^M = \Big\{(b_h, \{c_s, \nu_s, a_s\}_{s=0}^H) : \nu_0 \neq 0\Big\}$. In particular, the only parameter entering the denominator of g^M is a_0.

Assumption A.10. *Non-zero Impact effect of ε_{rt} in the Markov setting:*

$$a_0 \neq 0$$

This assumption requires that the instrument for the GE variable has a non-zero effect upon impact and implies that g^M is continuous at the true parameter vector.

For forward-looking settings, the portable elasticity at horizon h is defined as:

$$\beta_h^P = g^{FL}\Big(b_h, \{c_{hf}\}_{f=0}^F, \{\nu_s, \{a_{sf}\}_{f=0}^F\}_{s=0}^H\Big) \tag{A.101}$$

where g^{FL} is a continuous mapping on the domain $\Theta_h^{FL} = \Big\{(b_h, \{c_{hf}\}_{f=0}^F, \{\nu_s, \{a_{sf}\}_{f=0}^F\}_{s=0}^H) :$ $\boldsymbol{A}$ has full column rank$\}$. $\boldsymbol{A}$ is a matrix that collects the impulse responses of R to the news shocks across horizons.

Assumption A.11. *Rank Condition in the Forward-looking setting.*
The matrix $\boldsymbol{A}$ has full column rank.

This condition ensures the uniqueness of the solution to the least squares problem and guarantees that the mapping from time-series coefficients to the vector of counterfactual shocks, and hence to β_h^P, is continuous in a neighborhood of the true parameter vector. This is a sufficient condition for consistency.[6]

Since the parameter vector is finite-dimensional, marginal convergence in probability of each component implies joint convergence in probability. Given consistency of $(\hat{b}_h, \hat{c}_{hf}, \hat{v}_h, \hat{a}_h)$ for all $h \leq H$ and for all $f \in F$, it follows that

$$\hat{\Theta}_h^M = (\hat{b}_h, \{\hat{c}_s\}_{s=0}^H, \{\hat{v}_s, \hat{a}_s\}_{s=0}^H) \xrightarrow{p} (b_h, \{c_s\}_{s=0}^H, \{v_s, a_s\}_{s=0}^H) = \Theta_h^M, \tag{A.102}$$

$$\hat{\Theta}_h^{FL} = (\hat{b}_h, \{\hat{c}_{hf}\}_{f=0}^F, \{\hat{v}_s, \{\hat{a}_{sf}\}_{f=0}^F\}_{s=0}^H) \xrightarrow{p} (b_h, \{c_{hf}\}_{f=0}^F, \{v_s, \{a_{sf}\}_{f=0}^F\}_{s=0}^H) = \Theta_h^{FL}. \tag{A.103}$$

By the Continuous Mapping Theorem,

$$\hat{\beta}_h^P = g^k(\hat{\Theta}_h^k) \xrightarrow{p} g^k(\Theta_h^k) = \beta_h^P \quad \text{for } k = M, FL.$$

A.1.7 Inference

This appendix describes the procedure to compute standard errors for the estimated portable elasticities.

Let θ^h denote the $(K_p + K_t) \times 1$ vector of parameters at horizon h, where K_p is the number of panel parameters and K_t is the number of time-series parameters. The set of parameters $\Theta = [\theta^{h'}]_{h=1}^H$ can be estimated via just identified GMM. For panel regressions, assume the data are already two-way demeaned; let $\tilde{x}_{it}$ represent the two-way demeaned version of the variable x_{it}. The moment conditions for each horizon h are given by:

$$m_k^{P,h}(\theta^h) = \frac{1}{NT} \sum_{i,t} \tilde{x}_{it}^k \tilde{e}_{i,t+h}, \quad k = 1, \ldots, K_p, \tag{A.104}$$

$$m_{k'}^{T,h}(\theta^h) = \frac{1}{T} \sum_{t} x_t^{k'} e_{t+h}, \quad k' = 1, \ldots, K_t, \tag{A.105}$$

where $\tilde{e}_{i,t+h}$ and e_{t+h} represent residuals from the panel and time-series regressions, respectively. Throughout, I refer to the first set as *panel moment conditions* and the second set as

[6] Full column rank of A is a sufficient but not necessary condition for consistency. If A is rank deficient, the least squares problem admits multiple solutions for the vector of counterfactual shocks. However, β_h^P remains uniquely defined provided that the response of the unit-level outcome to the news shocks lies in the row space of A. Equivalently, for any vector d in the null space of A, it must hold that $\sum_{f=0}^F c_{hf} d_f = 0$ for all h. Economically, this condition requires that any linear combination of news shocks that leaves the path of R unchanged must also leave the path of the unit-level outcome unchanged. This is implied by the assumption of instrument validity, which requires that the news shocks affect the unit-level outcome only through their equilibrium effect on the path of R. Under this restriction, β_h^P depends only on the projection of $\{c_{hf}\}_f$ onto the row space of A, and full rank is not necessary for consistency.

time-series moment conditions. Define the full vector of moment conditions for horizon h as:

$$m^h(\theta^h) = \left(m^{P,h}(\theta^h)\ m^{T,h}(\theta^h)\right), \tag{A.106}$$

where $m^{P,h}(\theta^h)$ and $m^{T,h}(\theta^h)$ stack the panel and time-series moments, respectively. Let the stacked moment vector across all horizons be denoted as:

$$m(\Theta) = \left(m^1(\theta^1)\ m^2(\theta^2) \vdots m^H(\theta^H).\right) \tag{A.107}$$

The vector of parameters Θ can be estimated in one step by minimizing the quadratic form:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}}\, m(\Theta)' W m(\Theta), \tag{A.108}$$

where $W = I$, the identity matrix.

Standard Errors

The $(K_p + K_t)H \times (K_p + K_t)H$ covariance matrix of the moment conditions, $\hat{S}$, is calculated as follows. First, I assume zero covariance between moment conditions on different horizons h, implying that $\hat{S}$ is a block diagonal with blocks $\hat{S}^h$.[7] Within each horizon h:

1. For panel moments, the approach can flexibly accommodate three clustering approaches: (i) time, (ii) unit, and (iii) two-way clustering, denoted respectively by $\hat{S}^h_t$, $\hat{S}^h_i$, and $\hat{S}^h_{2w}$.
2. For time-series moments, I employ a Newey–West heteroskedasticity-and-autocorrelation consistent (HAC) estimator:

$$\hat{S}^h_{HAC} = \hat{\Gamma}(0) + \sum_{\ell=1}^{L} w(\ell)\left[\hat{\Gamma}(\ell) + \hat{\Gamma}(\ell)'\right], \tag{A.109}$$

where

$$\hat{\Gamma}(\ell) = \frac{1}{T}\sum_{t=\ell+1}^{T} m_t^{T,h}(\hat{\theta}^h) m_{t-\ell}^{T,h}(\hat{\theta}^h)', \quad w(\ell) = 1 - \frac{\ell}{L+1}. \tag{A.110}$$

[7] As an alternative I allow for covariance of moment conditions across horizons within each block (i.e., panel and time series moment conidtions). In that case, the covariance matrix is no longer block diagonal, and the off-diagonal blocks are estimated as

$$\hat{S}^{hh'} = \frac{1}{T}\sum_{t=1}^{T} m_t^h(\hat{\theta}^h)\, m_t^{h'}(\hat{\theta}^{h'})', \quad h \neq h'.$$

In Monte Carlo simulations, this approach leads to systematically larger standard errors and overcoverage when ignoring the cross-terms between panel and time series blocks. This may reflect the fact that moment conditions at different horizons load on the same underlying shocks, inducing strong cross-horizon dependence. For this reason, I adopt the block-diagonal specification as the baseline.

3. To capture the covariance between panel and time-series moments within the same time period, I define the aggregated moment vector at time t as

$$m_t^h(\theta^h) = \begin{pmatrix} m_t^{P,h}(\theta^h) \\ m_t^{T,h}(\theta^h) \end{pmatrix}, \tag{A.111}$$

where

$$m_t^{P,h}(\theta^h) = \begin{pmatrix} \frac{1}{N}\sum_{i=1}^N \tilde{x}_{i,t}^1 \tilde{e}_{i,t+h} \\ \vdots \\ \frac{1}{N}\sum_{i=1}^N \tilde{x}_{i,t}^{K_P} \tilde{e}_{i,t+h} \end{pmatrix}, \qquad m_t^{T,h}(\theta^h) = \begin{pmatrix} x_t^1 e_{t+h} \\ \vdots \\ x_t^{K_t} e_{t+h} \end{pmatrix}. \tag{A.112}$$

The time-clustered covariance between panel and time-series moments is then

$$\hat{S}_{PT}^h = \frac{1}{T}\sum_{t=1}^T m_t^{P,h}(\hat{\theta}^h) m_t^{T,h}(\hat{\theta}^h)'. \tag{A.113}$$

The covariance matrix for horizon h is given by

$$\hat{S}^h = \begin{pmatrix} \hat{S}_t^h & \hat{S}_{PT}^h \\ (\hat{S}_{PT}^h)' & \hat{S}_{HAC}^h \end{pmatrix}. \tag{A.114}$$

The full covariance matrix is block diagonal across horizons:

$$\hat{S} = \mathrm{diag}(\hat{S}^1, \hat{S}^2, \ldots, \hat{S}^H). \tag{A.115}$$

For completeness, I also consider allowing for covariance across horizons. For $h \neq h'$, define the off-diagonal block of the covariance matrix as

$$\hat{S}^{hh'} = \begin{pmatrix} \hat{S}_{PP}^{hh'} & \hat{S}_{PT}^{hh'} \\ \hat{S}_{TP}^{hh'} & \hat{S}_{TT}^{hh'} \end{pmatrix}, \tag{A.116}$$

where

$$\hat{S}_{PP}^{hh'} = \frac{1}{T}\sum_{t=1}^{T} m_t^{P,h}(\hat{\theta}^h)\, m_t^{P,h'}(\hat{\theta}^{h'})', \tag{A.117}$$

$$\hat{S}_{TT}^{hh'} = \frac{1}{T}\sum_{t=1}^{T} m_t^{T,h}(\hat{\theta}^h)\, m_t^{T,h'}(\hat{\theta}^{h'})', \tag{A.118}$$

$$\hat{S}_{PT}^{hh'} = \frac{1}{T}\sum_{t=1}^{T} m_t^{P,h}(\hat{\theta}^h)\, m_t^{T,h'}(\hat{\theta}^{h'})', \tag{A.119}$$

$$\hat{S}_{TP}^{hh'} = \frac{1}{T}\sum_{t=1}^{T} m_t^{T,h}(\hat{\theta}^h)\, m_t^{P,h'}(\hat{\theta}^{h'})'. \tag{A.120}$$

Under this specification, the full covariance matrix $\hat{S}$ is no longer block diagonal across horizons. I also consider a restricted version that rules out covariance between panel and time-series moments, both within and across horizons. In that case, I set

$$\hat{S}_{PT}^{h} = 0 \quad \text{for all } h, \qquad \hat{S}_{PT}^{hh'} = \hat{S}_{TP}^{hh'} = 0 \quad \text{for all } h \neq h', \tag{A.121}$$

while allowing $\hat{S}_{PP}^{hh'}$ and $\hat{S}_{TT}^{hh'}$ to be nonzero. Thus, the off-diagonal blocks reduce to

$$\hat{S}^{hh'} = \begin{pmatrix} \hat{S}_{PP}^{hh'} & 0 \\ 0 & \hat{S}_{TT}^{hh'} \end{pmatrix}, \qquad h \neq h'. \tag{A.122}$$

The asymptotic variance-covariance matrix of $\hat{\Theta}$ can be computed as:

$$\mathbb{V}(\hat{\Theta}) = \frac{1}{T}(G^\top \hat{S}^{-1} G)^{-1},$$

where G is the Jacobian of the moment conditions evaluated at $\hat{\Theta}$. The standard errors for the vector $\hat{\beta}^P = \{\hat{\beta}_h^P\}_{h=0}^H$ can be computed using the Delta method as follows:

$$\text{SE}(\hat{\beta}^P) = \sqrt{\nabla f^\top \mathbb{V}(\hat{\Theta}) \nabla f},$$

where ∇f is the Jacobian of β^P with respect to the vector of estimated parameters.

To evaluate the properties of the proposed inference procedure, I conduct Monte Carlo simulations in both static and dynamic settings.

In the static case, standard errors obtained from the joint GMM system coincide with those from the control function approach only when the covariance between panel and time-series moment conditions is properly accounted for. Ignoring this covariance leads to substantial overcoverage. Thus, the equivalence result for point estimates between the decomposition and control function approaches extends to inference when the shared variation across panel and time-series moments is incorporated.

In the dynamic setting, I consider four alternative estimators of the covariance matrix $\hat{S}$, which differ in whether they allow for (i) covariance between panel and time-series moments within a horizon ("cross-moments") and (ii) covariance across horizons ("cross-h"). The four cases are: (i) allowing cross-moments but imposing independence across horizons; (ii) imposing independence both across blocks and across horizons; (iii) allowing covariance both across blocks and across horizons; and (iv) allowing covariance across horizons within blocks while imposing zero covariance between panel and time-series moments.

The simulation results in Figure A.6 show that accounting for covariance across blocks, as in the static setting, improves coverage for the impact elasticity, but this does not extend to longer horizons. Beyond impact, allowing for cross-moment covariance leads to systematic undercoverage, regardless of whether cross-horizon dependence is included. In contrast, allowing for cross-horizon covariance alone leads to overcoverage. For this reason, I adopt as a baseline the specification that sets both cross-moment and cross-horizon covariances to zero, which displays the right coverage.

In addition, I conduct a Monte Carlo to explore coverage rates in small samples. Concretely, I repeat the exercise setting $N = 50$ and $T = 40$, similar to the sample size of the empirical application in Section 1.4. Figure A.7 presents the results. The baseline specification (ii) exhibits some undercoverage at longer horizons, but still delivers the most stable coverage profile among the estimators considered.

Lastly, Figure A.8 presents simulation results using the baseline specification (ii) replacing time-clustering by unit-clustering for panel moments. This exercise sets the sample back to $N = 100$ and $T = 300$. The results are in line with the findings in Almuzara and Sancibrián (2024), who point out that clustering at the unit level with common aggregate shocks can lead to substantial undercoverage.

Standard Error for Estimated Exante Counterfactual Shock

I briefly outline the construction of the standard error of $\boldsymbol{\varepsilon}_r^t$ for the single-shock case in the ex ante decomposition. The standard error of the ex ante shock is computed using the Delta method, treating the estimated impulse responses as jointly estimated. The ex ante shock size $\boldsymbol{\varepsilon}_r^t$ is defined as the scalar that minimizes the weighted distance between $\hat{\boldsymbol{R}}_g$ and $\hat{\boldsymbol{R}}_r$,

$$\boldsymbol{\varepsilon}_r^t = \arg\min_{\varepsilon} (\hat{\boldsymbol{R}}_g - \varepsilon \hat{\boldsymbol{R}}_r)' W (\hat{\boldsymbol{R}}_g - \varepsilon \hat{\boldsymbol{R}}_r) = \frac{\hat{\boldsymbol{R}}_r' W \hat{\boldsymbol{R}}_g}{\hat{\boldsymbol{R}}_r' W \hat{\boldsymbol{R}}_r}. \tag{A.123}$$

As a baseline, I set $W = 1/v(\hat{\boldsymbol{R}}_g)$. Let $z = (\hat{\boldsymbol{R}}_g', \hat{\boldsymbol{R}}_r')'$ collect the estimated impulse responses, with covariance matrix $\hat{S}_z$ obtained from the joint GMM system described above. The asymptotic variance of $\hat{\boldsymbol{\varepsilon}_r}^t$ is then given by

$$\mathbb{S}(\hat{\boldsymbol{\varepsilon}}_r^t) = \nabla_z \boldsymbol{\varepsilon}_r^{t\prime} \hat{S}_z \nabla_z \boldsymbol{\varepsilon}_r^t, \tag{A.124}$$

where $\nabla_z \boldsymbol{\varepsilon}_r^t$ denotes the gradient of $\boldsymbol{\varepsilon}_r^t$ with respect to z. This approach accounts for estimation uncertainty in the estimated impulse responses of R to each aggregate shock.

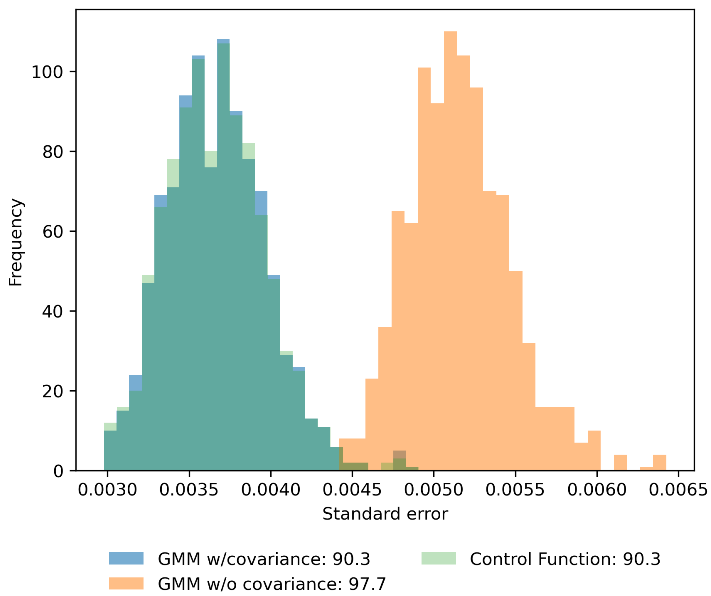


Figure A.5. Inference - MC Simulation - Static Setting.

Note: The plot shows the distribution of standard errors from the control function approach (light green), GMM with time-cluster covariance between panel and time series moments (blue) and GMM with zero covariance across panel and time series moments. The coverage rate is shown in the legend next to the label for each case. Based on 1000 repetitions.

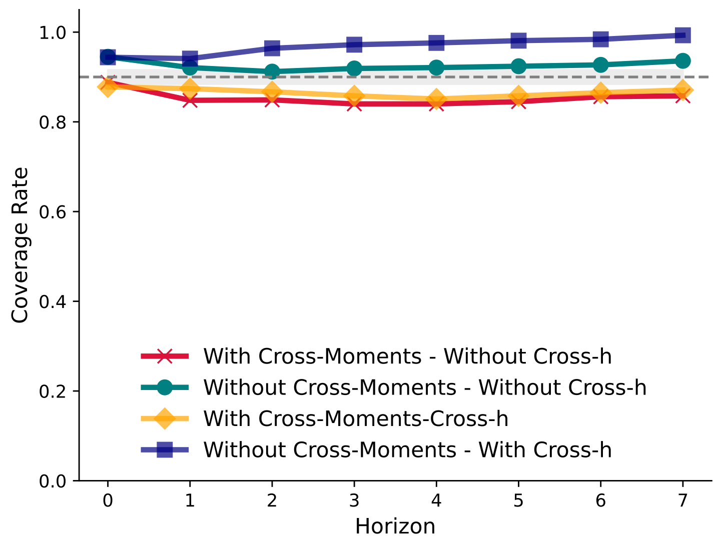


Figure A.6. Inference - MC Simulation - Dynamic Setting.

Note: The figure reports coverage rates across horizons h for four alternative estimators of the covariance matrix $\hat{S}$. *With Cross-Moments – Without Cross-h*: $\hat{S}$ allows for covariance between panel and time-series moment conditions within each horizon (time clustering), but remains block diagonal across horizons. *Without Cross-Moments – Without Cross-h*: $\hat{S}$ is block diagonal both across horizons and between panel and time-series moments. *With Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance both across horizons and between panel and time-series moment conditions. *Without Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance across horizons within panel and within time-series moments, while imposing zero covariance between panel and time-series blocks. All results are based on 1000 Monte Carlo repetitions with N=100 and T=300. The DGP is parametrized as $\rho_g = \rho_r = .8, \delta = .5, \beta = 1, \gamma = .5, \phi = .5$. The variance of all innovations is set to unity.

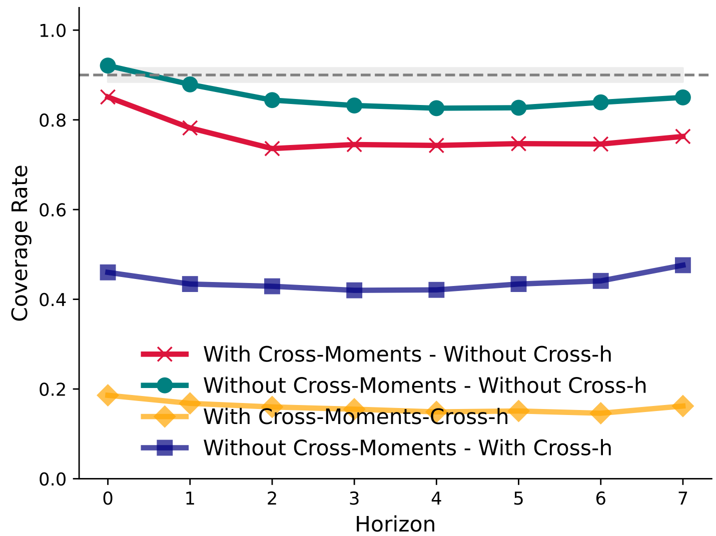


Figure A.7. Inference - MC Simulation - Dynamic Setting with Small Sample.

Note: The figure reports coverage rates across horizons h for four alternative estimators of the covariance matrix $\hat{S}$. *With Cross-Moments – Without Cross-h*: $\hat{S}$ allows for covariance between panel and time-series moment conditions within each horizon (time clustering), but remains block diagonal across horizons. *Without Cross-Moments – Without Cross-h*: $\hat{S}$ is block diagonal both across horizons and between panel and time-series moments. *With Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance both across horizons and between panel and time-series moment conditions. *Without Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance across horizons within panel and within time-series moments, while imposing zero covariance between panel and time-series blocks. All results are based on 1000 Monte Carlo repetitions with $N = 50$ and $T = 40$. The DGP is parametrized as $\rho_g = \rho_r = .8, \delta = .5, \beta = 1, \gamma = .5, \phi = .5$. The variance of all innovations is set to unity.

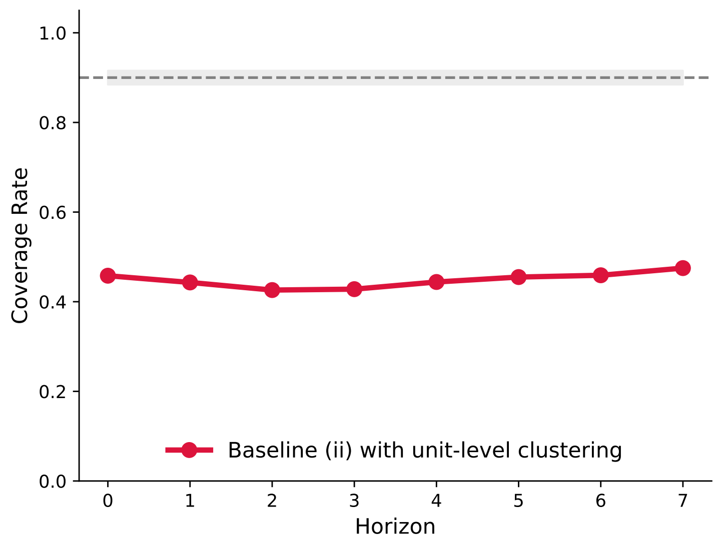


Figure A.8. Inference - MC Simulation Dynamic Setting with Unit-level Clustering.

Note: The figure reports coverage rates across horizons h for the *Without Cross-Moments – Without Cross-h* estimator of the covariance matrix $\hat{S}$ using unit-level clustering for panel moment conditions. All results are based on $J = 1000$ Monte Carlo repetitions with $N = 100$ and $T = 300$. The DGP is parametrized as $\rho_g = \rho_r = .8, \delta = .5, \beta = 1, \gamma = .5, \phi = .5$. The variance of all innovations is set to unity.

A.2 Data

This appendix presents additional details and descriptive statistics on the data used for the estimation of cross-sectional fiscal multipliers in Section 1.4.

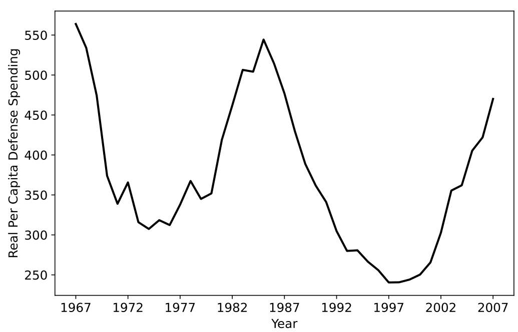
In Table A.6 I present descriptive statistics for the main time series and cross-sectional variables used in the regression analysis. The sample period is 1967-2007. G_t^C measures the one-year change in real per-capita defense contracts relative to lagged GDP. The average change in G_t^C is close to zero with a standard deviation of around .3 percent of output. The standard deviation of the Romer-Romer monetary policy shock is close to 1 percentage point. The average exposure to defense spending is .85 with substantial heterogeneity across space. The participation of defense spending in output for the most exposed region is more than three times that of the aggregate economy, whereas for the least exposed one the participation is close to one tenth.

Table A.6. Descriptive Statistics for Time Series and Cross-Sectional Variables

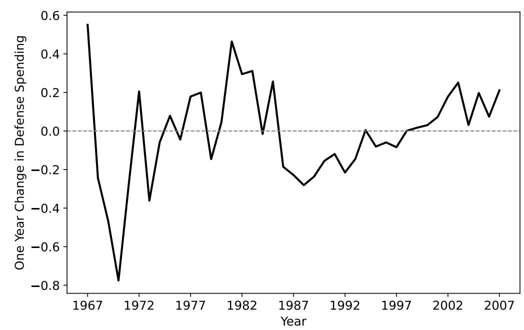
Variable	Mean	Median	Std Dev	Min	Max	Units
Aggregate Variables						
G_t^C	-0.01	-0.04	0.28	-0.78	1.03	% of Y_{t-1}^{pc}
T-bill Rate (3M)	4.20	4.25	3.02	0.03	14.03	%
Real T-bill Rate (3M)	1.19	1.39	1.93	-3.37	5.45	%
Romer-Romer MP shock	0.00	0.21	0.93	-2.18	1.89	%
Cross-sectional Variables						
Y_{it}^C	1.68	1.65	3.82	-29.52	34.39	% of Y_{it-1}^{pc}
G_{it}^C	-0.01	-0.01	0.88	-9.04	7.58	% of Y_{it-1}^{pc}
s_{ig}^{pre}	0.85	0.79	0.54	0.15	3.19	
$\overline{Y}_{it}^{pc}$	16731	16067	3107	11581	30096	Real p.c.

Table A.7. Top 5 States by Exposure to Defense Spending - 1966-70 Average

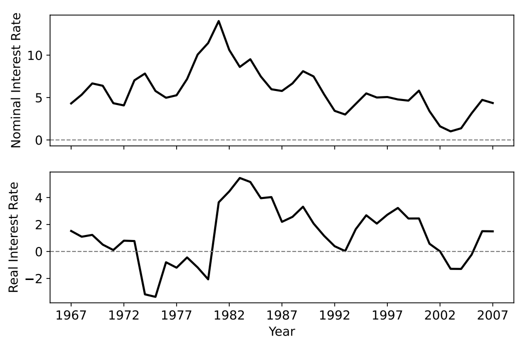
Top 5 States - 1966-70 Average	s_{ig}^{pre}
Connecticut	3.19
Texas	1.87
Missouri	1.70
California	1.68
Massachusetts	1.48



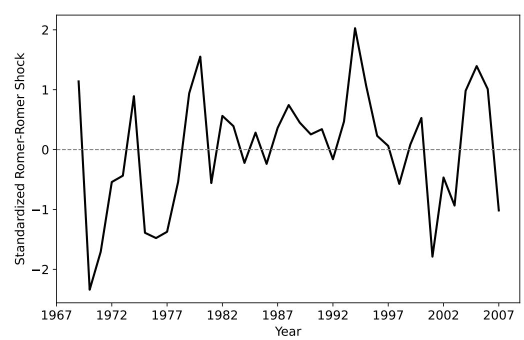
(A) National Defense Spending (p.c.).



(B) Δ_1 National Defense Spending (% of GDP).



(C) 3-month T-bill Rate (%).



(D) Romer-Romer Monetary Shock Series.

Figure A.9. Time Series Data.

Note: Nominal variables are deflated using the consumer price index. The Romer-Romer monetary shock series is expressed in standard deviations around a zero mean.

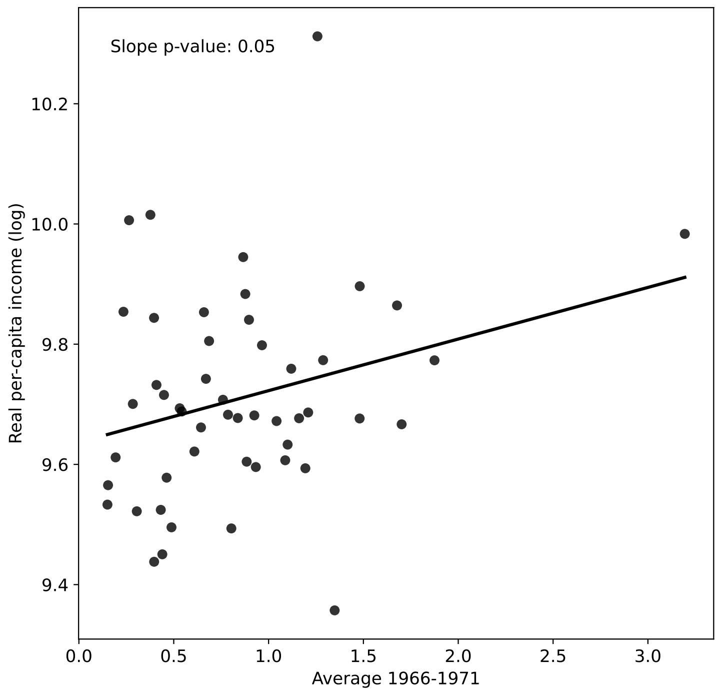


Figure A.10. Exposure to Defense Spending and Per-Capita Income.

Note: Nominal variables are deflated using the consumer price index.

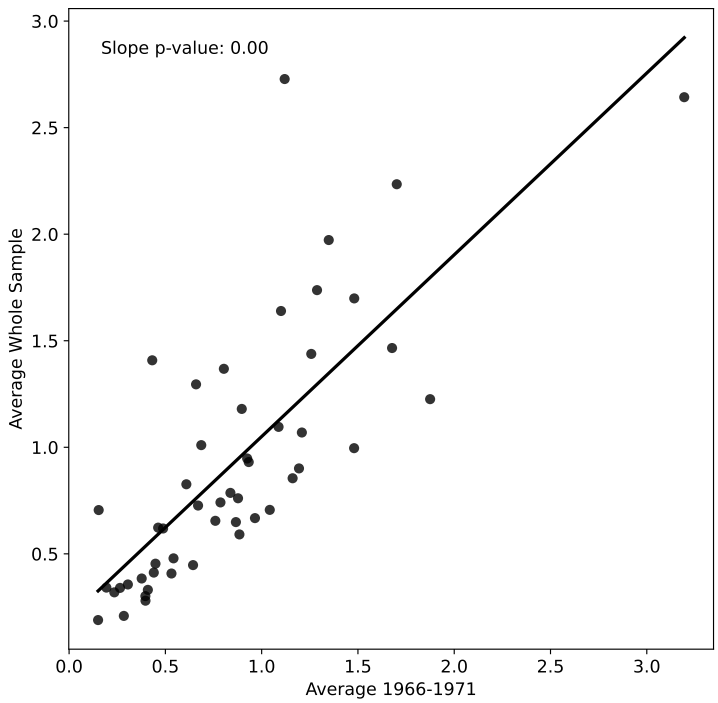


Figure A.11. Pre-period and Period-Average Exposure to Defense Spending.

Note: Pre-period corresponds to the average for 1966-1970 while Period-average corresponds to the average for 1969-2007.

A.3 Estimation Results

A.3.1 HEDGE Test - Robustness

Figure A.12 presents results when including the interaction between s_{ig} and the fiscal shock. Formally:

$$Y^C_{it+h} = b_h \times s_{ig} \times G^C_t c_h \times s_{ig} i_t + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h} \qquad \text{for} \quad h = 1,4. \tag{A.125}$$

The estimated responses are quantitatively unchanged relative to the baseline estimates in Figure 1.4.

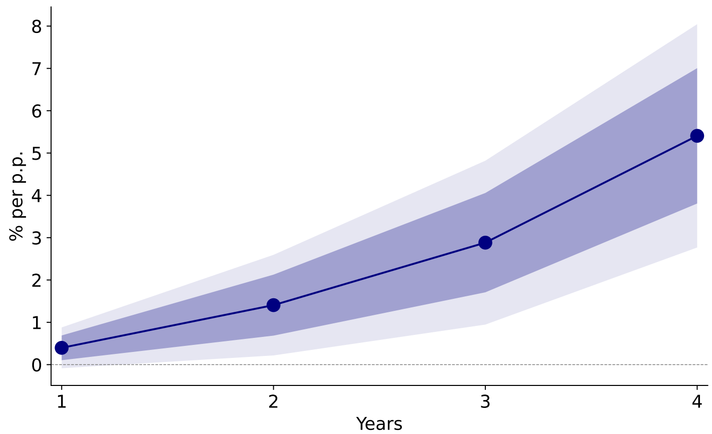


Figure A.12. HEDGE Test - Robustness to Inclusion of $s_{ig} \times G^C_t$.

Note: Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. The sample covers 1970-2007. The point estimates measure the cumulative percent change in local output per percentage point increase in the Federal Funds rate.

Figure A.13 plots the estimated responses when using the series of shocks of Aruoba and Drechsel (2024) instead of the Romer and Romer (2004) shocks. The sample is reduced to the years between 1980-2007 because of the shock series availability.

Lastly, I estimate state-specific output elasticities to changes in the interest rate. For each state j and horizon h, I estimate:

$$Y^C_{jt+h} = \sum_{s=0}^{S} c^j_h \times RR_{t-s} + \phi^j_h Y^C_{jt-1} + e_{jt+h}, \tag{A.126}$$

where c^j_h measures the response of Y^C_{jt+h} to a one standard deviation Romer and Romer (2004) monetary shock. Figure A.14 plots the estimated responses against s_{ig} - the exposure to defense spending used for identification. Each panel corresponds to a different horizon and the solid black line shows the estimated linear fit. Estimate state-specific elasticities are positively correlated with s_{ig} at all four horizons. The estimated linear relationship is statistically significant at least at the 5% level for all horizons.

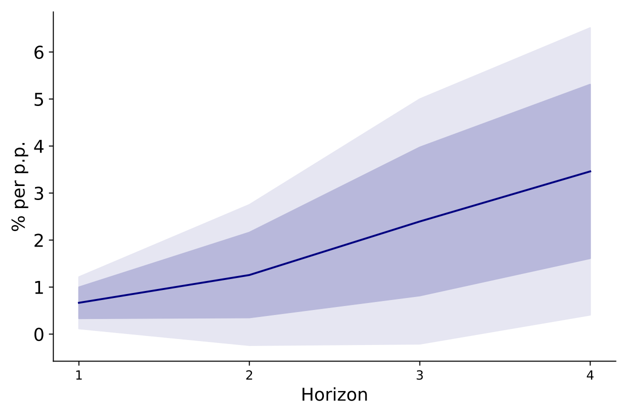
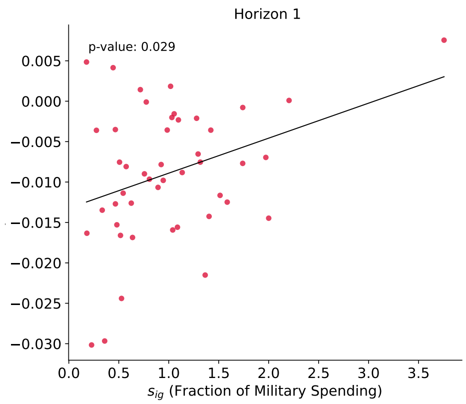
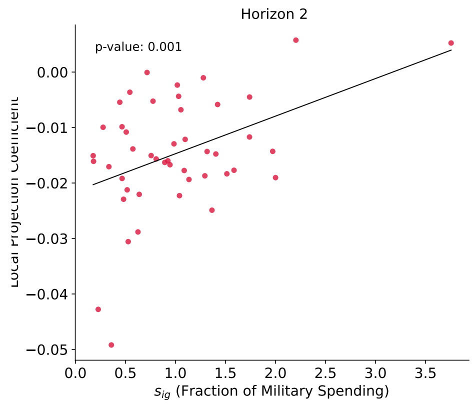


Figure A.13. HEDGE Test- Differential Response to i_t - Aruoba and Drechsel (2024).

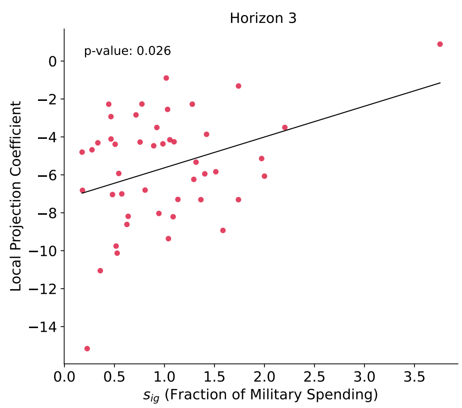
Note: Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. The sample covers 1980-2007. The point estimates measure the cumulative percent change in local output per percentage point increase in the Federal Funds rate.



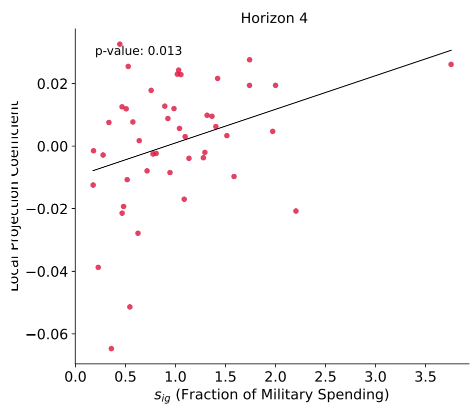
(A) $h = 1$.



(B) $h = 2$.



(C) $h = 3$.



(D) $h = 4$.

Figure A.14. State-Specific Interest Rate Elasticities and s_{ig}.

A.3.2 Decomposition of the US Cross Sectional Multiplier - Additional Results

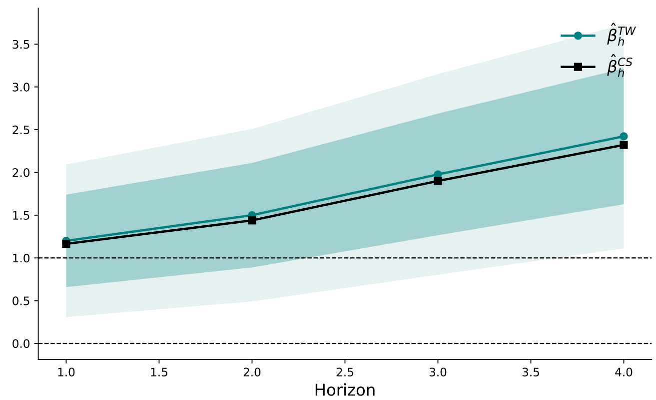


Figure A.15. Comparison between $\hat{\beta}_h^{TW}$ and $\hat{\beta}_h^{CS}$.

Note: Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. The superscript TW denotes the TWFE specification and CS denotes the estimate from the cross-sectional step of the decomposition framework. Both specifications control for lagged output growth. Time-clustered standard errors.

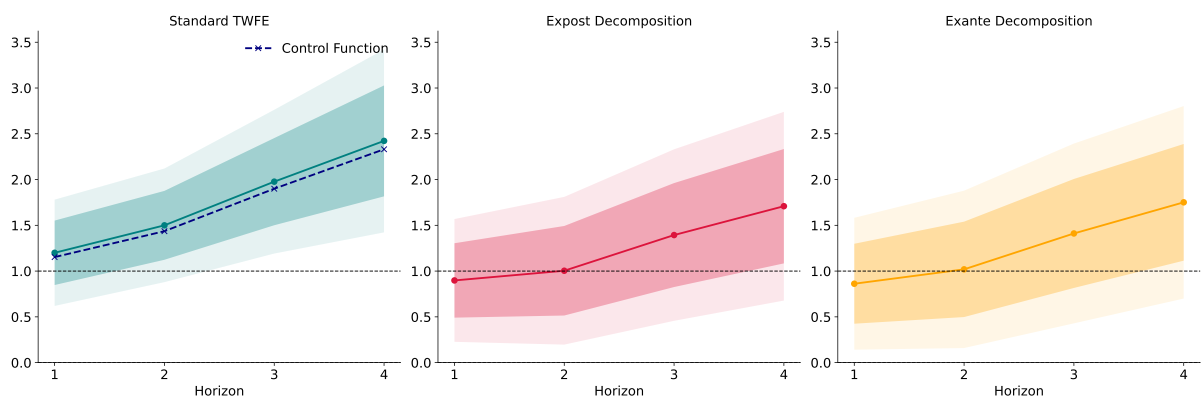


Figure A.16. Decomposition Results for the US Multiplier - Unit-level Clustering.

Note: Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.1.7 with unit-level clustering for panel moment conditions. Point estimates show the cumulative percentage change in local output when local defense spending increases by 1% of local output. *Ex-post Decomposition* shows results for the decomposition outlined in Subsection 1.3.2. *Ex-ante Decomposition* shows results for the decomposition outlined in Subsection 1.3.3 using a single shock. The blue dotted line shows the results from the control function approach.

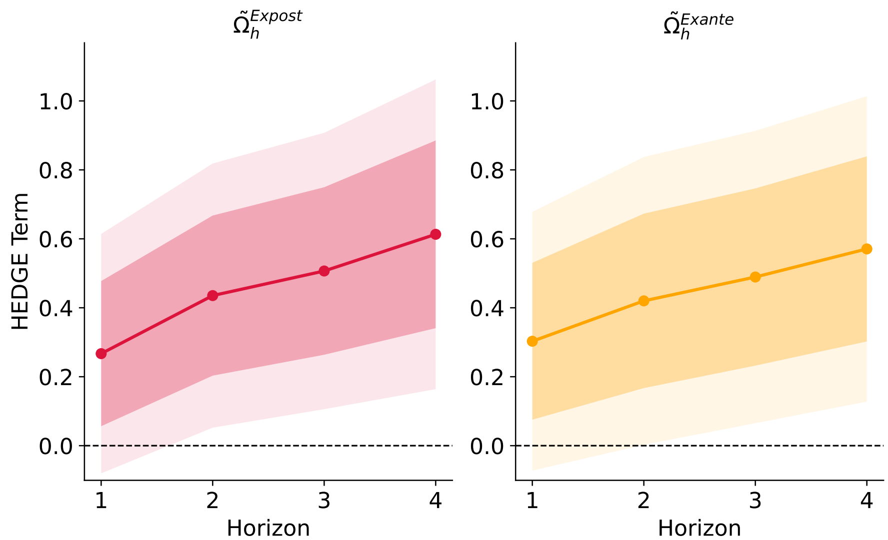


Figure A.17. Estimated HEDGE Terms - Unit-level Clustering.

Note: The plot shows the point estimate and confidence bands for $\tilde{\Omega}_{h,Expost}$ on the left-hand side and for $\tilde{\Omega}_{h,Exante}$ on the right-hand side. Standard errors are computed as outlined in Subsection 1.3.4 with unit-level clustering for panel moment conditions.

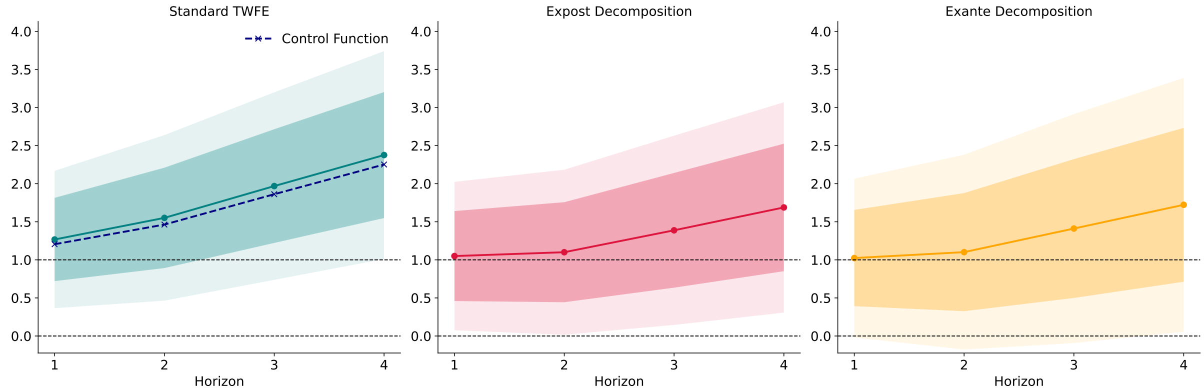
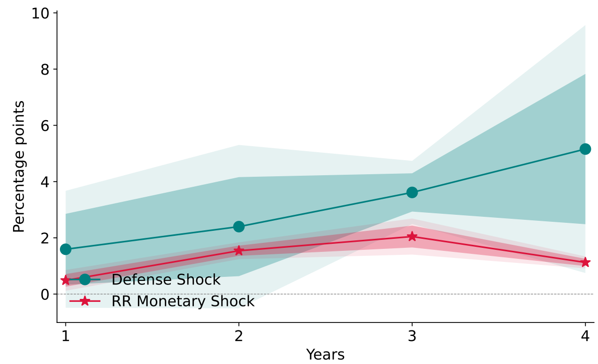
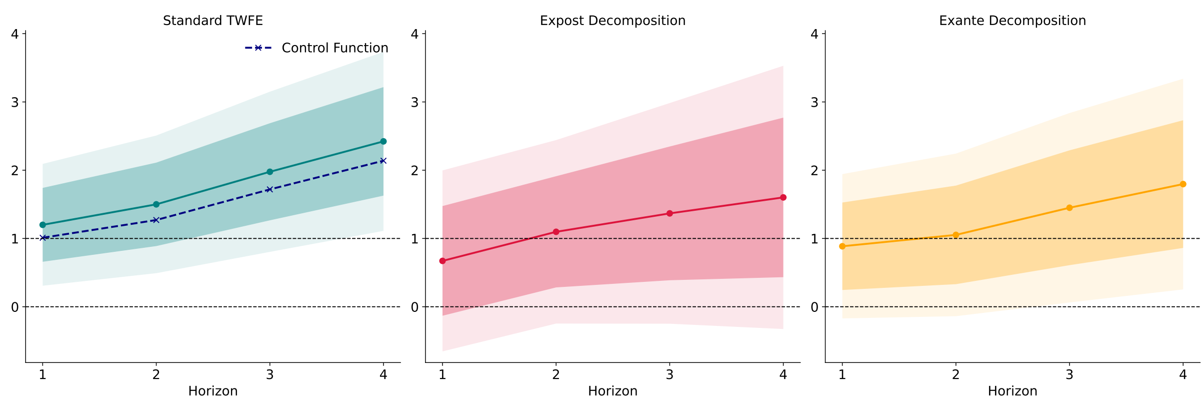


Figure A.18. Decomposition Results for the US Multiplier - All States.

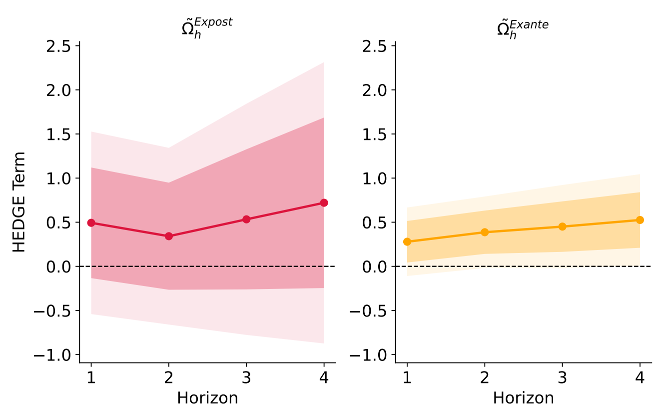
Note: Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.1.7 with time-level clustering for panel moment conditions. Point estimates show the cumulative percentage change in local output when local defense spending increases by 1% of local output. *Ex-post Decomposition* shows results for the decomposition outlined in Subsection 1.3.2. *Ex-ante Decomposition* shows results for the decomposition outlined in Subsection 1.3.3 using a single shock. The blue dotted line shows the results from the control function approach.



(A) Response of Real Federal Funds Rate.



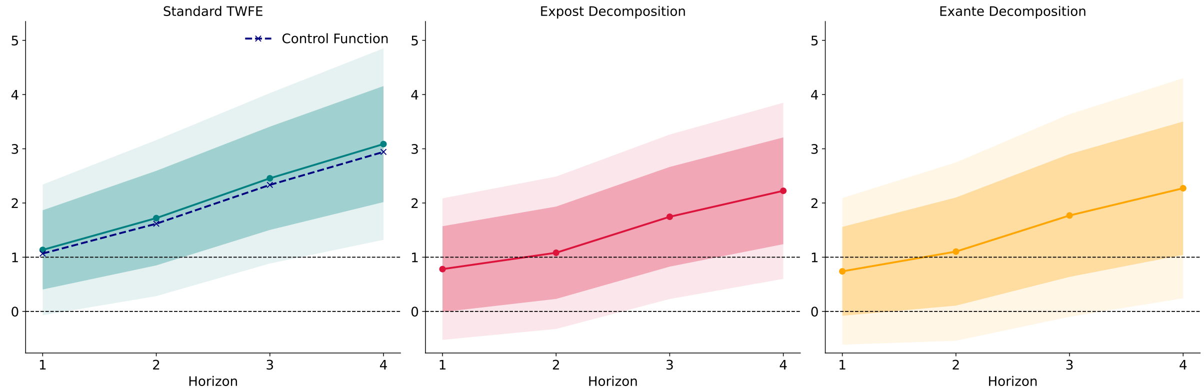
(B) Cross-sectional Fiscal Multipliers.



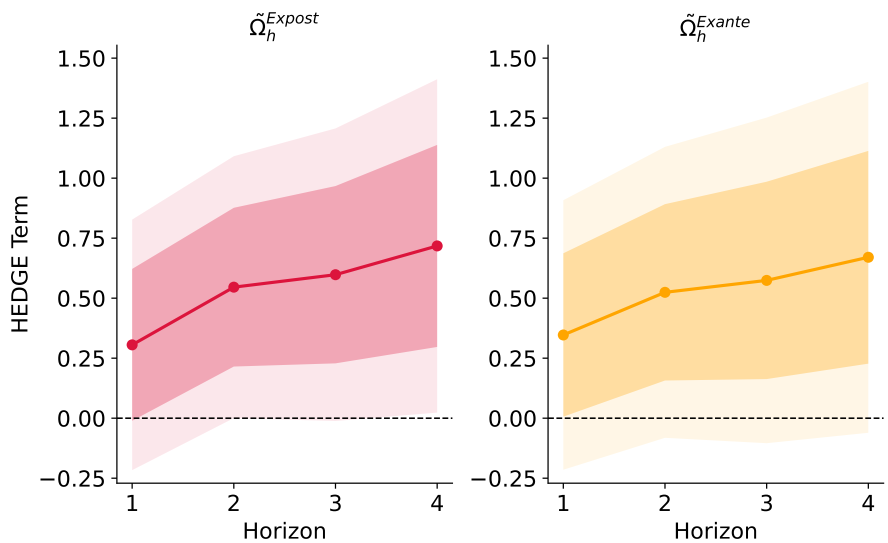
(C) HEDGE Terms.

Figure A.19. Results for the Real Federal Funds Rate.

Note: Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results from using the estimated path of the real Federal Funds rate to construct ex-post and ex-ante sequences of counterfactual shocks. All specifications include the same set of controls as in the baseline. Standard errors are computed as outlined in Subsection 1.3.4 with time-level clustering for panel moment conditions.



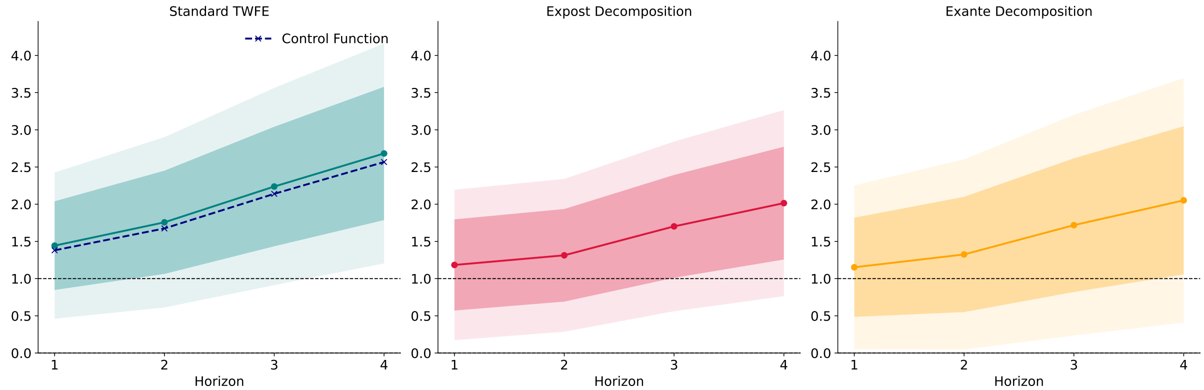
(A) Cross-sectional Fiscal Multipliers.



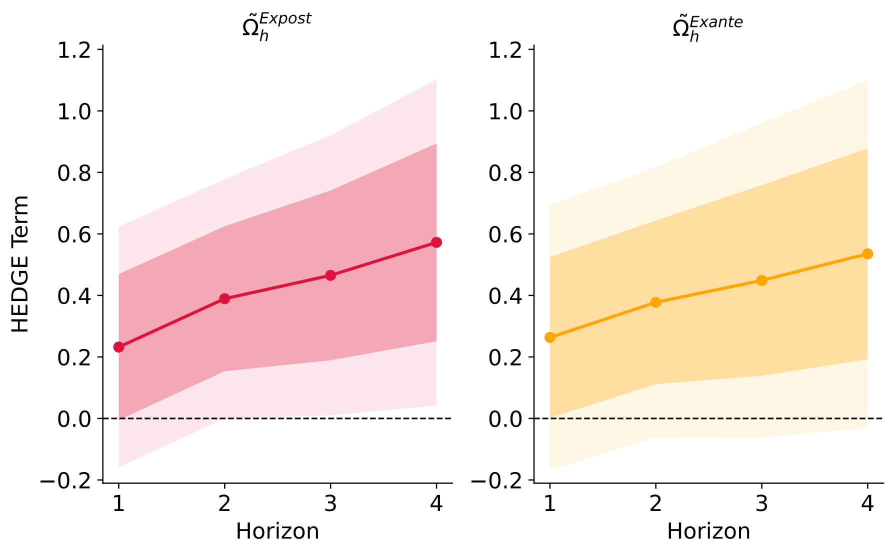
(B) HEDGE Terms.

Figure A.20. Average Exposure Share 1966-2007.

Note: Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results of using the average exposure to defense spending between 1966-2007 instead of the average during the pre-period. Standard errors are computed as outlined in Subsection 1.3.4 with time-level clustering for panel moment conditions.



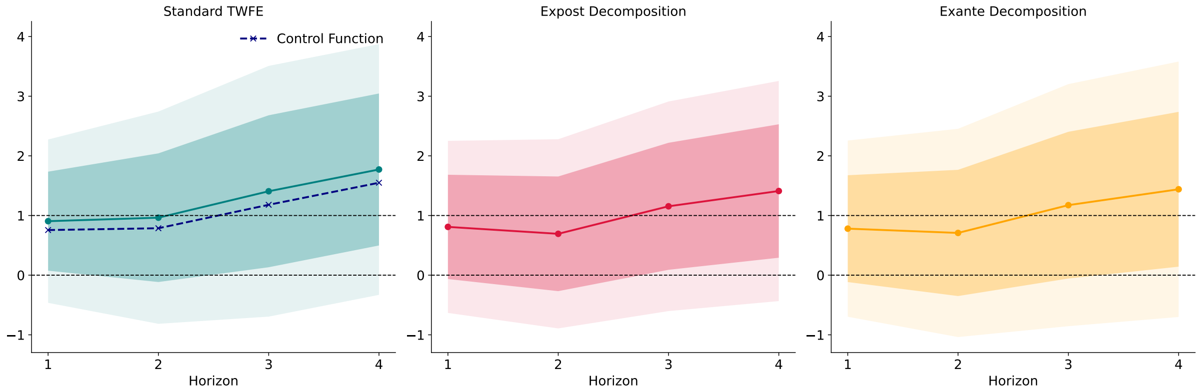
(A) Cross-sectional Fiscal Multipliers.



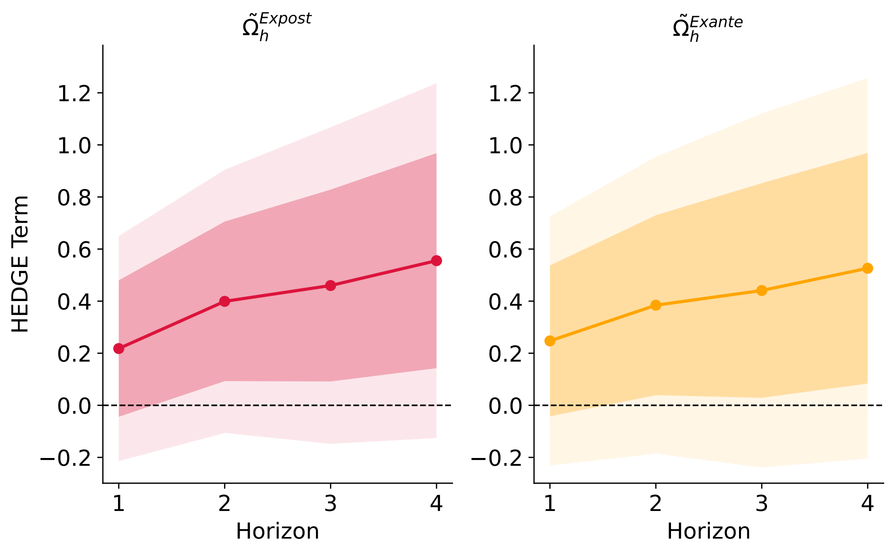
(B) HEDGE Terms.

Figure A.21. No Controls in Cross-sectional Regressions.

Note: Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results from dropping the control for lagged output growth. Standard errors are computed as outlined in Subsection 1.3.4 with time-level clustering for panel moment conditions.



(A) Cross-sectional Fiscal Multipliers.



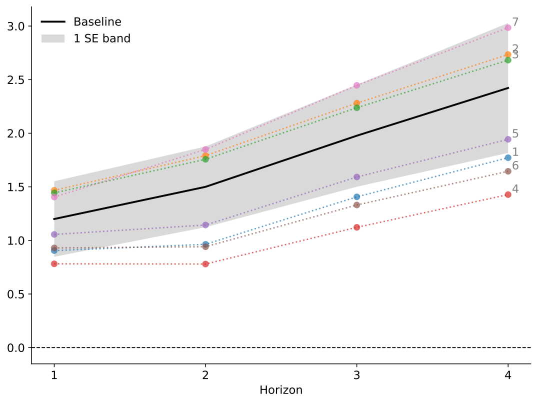
(B) HEDGE Terms.

Figure A.22. Adding Lagged Shocks as Controls.

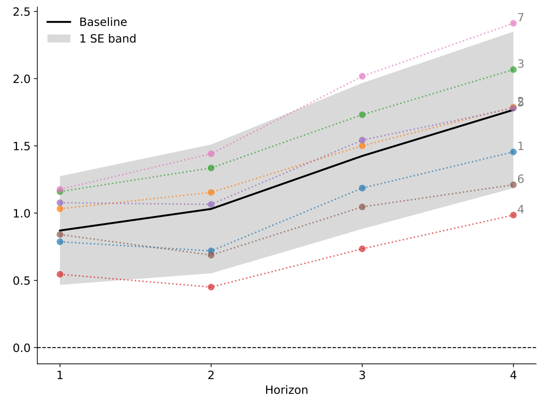
Note: Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results from adding $s_{ig}G^C_{t-1}$ as control in the TWFE regression and, both $s_{ig}G^C_{t-1}$ and $s_{ig}RR_{t-1}$ in the cross-sectional step regression. Standard errors are computed as outlined in Subsection 1.3.4 with time-level clustering for panel moment conditions.

Table A.8. Alternative Controls in Cross-sectional Regressions.

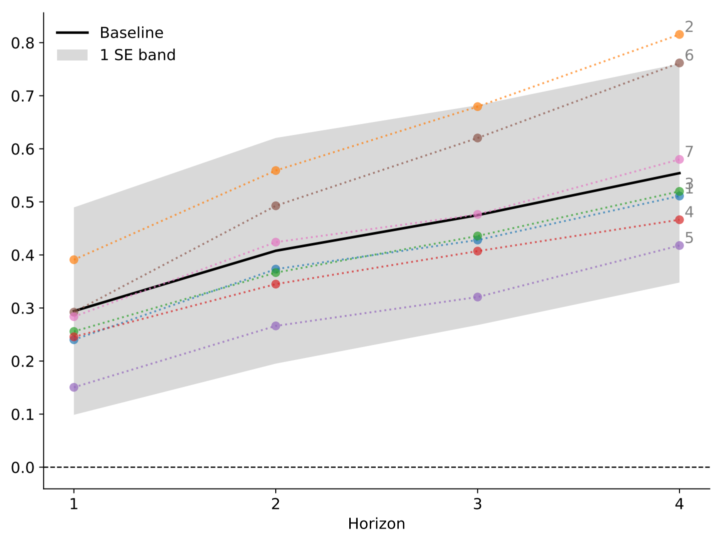
Spec #	**Included Controls**
Baseline	Y^C_{it-1}
1	Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$
2	$\ln(Y_{it-1})$
3	No controls
4	Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$
5	$s_{ig} \times RR_{t-1}$, $s_{ig} \times G^C_{t-1}$
6	$\ln(Y_{it-1})$, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$
7	Y^C_{it-1}, $s_{ig} \times RR_{t-1}$



(A) TWFE Multiplier.



(B) Portable Multipliers - Ex-ante Decomposition.



(C) HEDGE Term - Ex-ante Decomposition.

Figure A.23. Alternative Controls.

Note: Shaded areas correspond to a one standard deviation. Baseline: Y^C_{it-1}; 1: Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$; 2: $\ln(Y_{it-1})$; 3: No controls; 4: Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$; 5: $s_{ig} \times RR_{t-1}$, $s_{ig} \times G^C_{t-1}$; 6: $\ln(Y_{it-1})$, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$; 7: Y^C_{it-1}, $s_{ig} \times RR_{t-1}$.

A.4 Model Details

This Appendix presents a full model of a monetary union with two heterogeneous regions labeled Home and Foreign. The model is an extension of Nakamura and Steinsson (2014) to include two types of households - *Ricardians* and *Hand-to-Mouth* - as in Herreño and Pedemonte (2025). In addition, regions can differ in their (i) household composition, (ii) intertemporal elasticity of substitution, (iii) size, and (iv) exposure to government spending shocks.

A.4.1 Home Households

The total population of the monetary union is normalized to one. The size of the Home region is n, while the size of the Foreign region is $(1-n)$. The Home region is populated by two types of agents - *Ricardians* and *Hand-to-mouth* - indexed with subscripts R and C, respectively. The share of the Home population that is *Hand-to-mouth* is given by λ^H. Both households have the same preferences over labor and consumption, but differ in their access to asset markets. They maximize lifetime utility subject to their budget constraints.

$$\max_{C_{H,k,t+s},L_{H,k,t+s}} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^{t+s} U\Big(C_{H,k,t+s}, L_{H,k,t+s}\Big) \;\; \forall \;\; k = C, R. \tag{A.127}$$

The budget constraints for each type of household are the following:

$$\mathcal{P}_t C_{H,C,t} \leq (1-\tau_t) W_{H,t} L_{H,C,t} + \frac{(1-z)}{\lambda^H} \Upsilon_{H,t} - T_{C,t}, \tag{A.128}$$

$$\mathcal{P}_t C_{H,R,t} + B_{H,t+1} \leq (1-\tau_t) W_{H,t} L_{H,R,t} + (1+i_t) B_{H,t} + \frac{z}{1-\lambda^H} \Upsilon_{H,t} - T_{R,t}. \tag{A.129}$$

$\mathcal{P}_t$ is the Home CPI (i.e. the cost of acquiring one unit of the home consumption bundle). B_{t+1} are holdings of nominal bonds that pay the national interest rate i_{t+1}. $W_{H,t}$ is the nominal wage rate in the Home region. $L_{H,k,t}$ is per-capita labor supply of households of type k. $\Upsilon_{H,t}$ are aggregate nominal profits from firms in the Home region. A share z of the aggregate profits is transferred to the Ricardian households such that the profits per capita are given by $\frac{z}{1-\lambda}\Upsilon_{H,t}$, and similarly for the *hand-to-mouth* households. $T_{H,k,t}$ are lump-sum taxes levied by the national government on household type k.

The Home final consumption good, $C_{H,k,t}$, is a CES bundle of Home and Foreign composite goods, $C^H_{H,k,t}$ and $C^F_{H,k,t}$:

$$C_{H,k,t} = \left[\phi_H^{\frac{1}{\eta}} C^{H\;\frac{\eta-1}{\eta}}_{H,k,t} + (1-\phi_H)^{\frac{1}{\eta}} C^{F\;\frac{\eta-1}{\eta}}_{H,k,t} \right]^{\frac{\eta}{\eta-1}} \;\; \forall \;\; k = C, R. \tag{A.130}$$

I use superscripts to denote the region where production takes place and subscripts to denote the region where consumption takes place (i.e. $C^F_{H,k,t}$ is a composite of goods produced in region F consumed by household type k located in region H). The parameter η governs the elasticity of substitution between domestic and foreign goods. Parameter ϕ_H captures the share of

Home goods in total consumption and, hence, the extent of home bias. Both parameters govern the strength of expenditure-switching in response to changes in relative regional prices. Each regional composite good is a CES bundle of individual varieties.

$$C^H_{H,k,t} = \left[\int_0^1 c^H_{H,k,t}(z)^{\frac{\theta}{\theta-1}} dz\right]^{\frac{\theta-1}{\theta}}, \qquad C^F_{H,k,t} = \left[\int_0^1 c^F_{H,k,t}(z)^{\frac{\theta}{\theta-1}} dz\right]^{\frac{\theta-1}{\theta}}, \tag{A.131}$$

where $c^H_{H,k,t}(z)$ and $c^F_{H,k,t}(z)$ denote consumption of Home and Foreign variety z, respectively. The parameter θ governs the elasticity of substitution between varieties produced within a given region. Goods trade freely across regions; therefore, both regions face the same prices for each individual variety. I denote prices at the variety level by $p^H_t(z)$ and $p^F_t(z)$, respectively. Price indexes are defined as follows.

$$P_{H,t} = \left[\int_0^1 p_{H,t}(z)^{1-\theta} dz\right]^{\frac{1}{1-\theta}}, \qquad P_{F,t} = \left[\int_0^1 p_{F,t}(z)^{1-\theta} dz\right]^{\frac{1}{1-\theta}}, \tag{A.132}$$

and

$$\mathcal{P}_t = \left[\phi^H P^{1-\eta}_{H,t} + (1-\phi^H) P^{1-\eta}_{F,t}\right]^{\frac{1}{1-\eta}}. \tag{A.133}$$

$P_{H,t}$ and $P_{F,t}$ are the PPI prices indexes of goods produced in the Home and Foreign region, respectively. As mentioned above, $\mathcal{P}_t$ is the CPI price index of the Home region. I assume that preferences are separable in labor and consumption. Concretely,

$$U\Big(C_{H,k,t}, L_{H,k,t}\Big) = \frac{C^{1-\sigma_H{}^{-1}}_{H,k,t}}{1-\sigma_H{}^{-1}} - \psi \frac{L^{1+\frac{1}{\nu}}_{H,k,t}}{1+\frac{1}{\nu}}, \tag{A.134}$$

where σ_H is the intertemporal elasticity of substitution and ν is the Frisch elasticity of labor supply. Utility maximization by Ricardian households yields a standard Euler equation:

$$C^{-\sigma_H^{-1}}_{H,R,t} = \beta \mathbb{E}_t \left[C^{-\sigma_H^{-1}}_{H,R,t+1} (1+i_t) \frac{\mathcal{P}_t}{\mathcal{P}_{t+1}}\right]. \tag{A.135}$$

The intratemporal trade-off between labor and consumption is dictated by:

$$\frac{\psi L^{\nu^{-1}}_{H,k,t}}{C^{-\sigma_H{}^{-1}}_{H,k,t}} = (1-\tau_t) \frac{W_{H,t}}{\mathcal{P}_t} \quad k = C, R. \tag{A.136}$$

Note that this trade-off is the same for both Ricardians and *hand-to-mouth* households.

Lastly, minimization of costs to achieve a given consumption level $C_{H,k,t}$ implies the following demand schedules:

$$C^H_{H,k,t} = \phi^H C_{H,k,t}\left(\frac{P_{H,t}}{\mathcal{P}_t}\right)^{-\eta}, \qquad C^F_{H,k,t} = (1-\phi^H)C_{H,k,t}\left(\frac{P_{F,t}}{\mathcal{P}_t}\right)^{-\eta}, \tag{A.137}$$

$$c^H_{H,k,t}(z) = C^H_{H,k,t}\left(\frac{p_{H,t}(z)}{P_{H,t}}\right)^{-\theta}, \qquad c^F_{H,k,t}(z) = C^F_{H,k,t}\left(\frac{p_{F,t}(z)}{P_{F,t}}\right)^{-\theta}, \tag{A.138}$$

for $k = C, R$.

A.4.2 Foreign Households

The problem for the Foreign region is symmetric, so I only present some key equations here. The share of *Hand-to-mouth* households in the Foreign region is denoted by λ^F. The final good bundle of the Foreign region, $C_{F,k,t}$ is given by:

$$C_{F,k,t} = \left[\phi_F^{\frac{1}{\eta}} {C^F_{F,k,t}}^{\frac{\eta-1}{\eta}} + (1-\phi_F)^{\frac{1}{\eta}} {C^H_{F,k,t}}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \quad \forall \;\; k = C, R, \tag{A.139}$$

where ϕ_F governs the degree of home-bias in the Foreign region. The Foreign Euler equation for Ricardian households is:

$$C_{F,R,t}^{-\sigma_F^{-1}} = \beta\mathbb{E}_t\left[C_{F,R,t+1}^{-\sigma_F^{-1}}(1+i_t)\frac{\mathcal{P}^*_t}{\mathcal{P}^*_{t+1}}\right], \tag{A.140}$$

where $\mathcal{P}^*_t$ is the CPI price level in the Foreign region given by:

$$\mathcal{P}^*_t = \left[\phi^F P_{F,t}^{1-\eta} + (1-\phi^F)P_{H,t}^{1-\eta}\right]^{\frac{1}{1-\eta}}. \tag{A.141}$$

A.4.3 Risk-sharing Condition

Complete markets and the Euler equations of the Home and Foreign region imply the following risk-sharing condition:

$$\frac{C_{F,R,t}^{-\sigma_F^{-1}}}{C_{H,R,t}^{-\sigma_H^{-1}}} = \omega_t Q_t, \tag{A.142}$$

where $\omega_t = 1$ on a balanced growth path and $Q_t = \frac{\mathcal{P}^*_t}{\mathcal{P}_t}$ is the real exchange rate between regions in the monetary union. Note that the risk-sharing condition is in terms of Ricardian consumption only.

A.4.4 Household Aggregates

Per capita consumption and labor supply in the region r satisfy:

$$C_{r,t} = \lambda^r C_{r,C,t} + (1-\lambda^r) C_{r,R,t}, \qquad L_{r,t} = \lambda^r L_{r,C,t} + (1-\lambda^r) L_{r,R,t}, \tag{A.143}$$

for $r = H, F$. Aggregate consumption and labor supply satisfy:

$$C_t = nC_{H,t} + (1-n)C_{F,t}, \qquad L_t = nL_{H,t} + (1-n)L_{F,t}. \tag{A.144}$$

A.4.5 Production

The problem of Home and Foreign producers is symmetric, so I focus on Home firms. The economy features a continuum of monopolistically competitive producers of Home varieties, indexed by z, each operating the technology $y_t^H(z) = L_{H,t}^a(z)$. They take regional wages as given and set prices subject to a la Calvo stickiness. The demand schedule for firm z is given by:

$$y_t^H(z) = \left[nC_{H,t}^H + (1-n)C_{F,t}^H + nG_t^H \right] \left(\frac{p_{H,t}(z)}{P_{H,t}} \right)^{-\theta}, \tag{A.145}$$

where $C_{H,t}^H$ and $C_{F,t}^H$ are the per-capita demands for Home goods coming from Home and Foreign households, respectively. G_t^H is per-capita government demand of the Home good.[8] Firms take regional wages, $W_{H,t}$, as given such that nominal marginal costs are:

$$MC_t^H(z) = \frac{W_{H,t}}{aL_{H,t}^{a-1}(z)}. \tag{A.146}$$

Firms get to reset prices with probability $1-\alpha$ every period (Calvo, 1983). They choose prices to maximize the discounted value of future real profits subject to the demand schedule in (A.145). Concretely, the problem of a firm that gets to reset prices at t is:

$$\max_{p_{H,t}^*(z)} \sum_{s=0}^{\infty} \alpha^s \Lambda_{t+s|t} y_{t+s|t}^H(z) \left[p_{H,t}^*(z) - \frac{W_{H,t+s}}{aL_{H,t+s}^{a-1}(z)} \right] \qquad \text{s.t.} \quad (\text{A.145}), \tag{A.147}$$

where $\Lambda_{t+s|t}$ is the stochastic discount factor of the Ricardian household, $p_{H,t}^*(z)$ is the optimal reset price and $y_{t+s|t}^H(z)$ is the demand faced in period $t+s$ conditional on resetting prices in period t. The solution to this problem yields the following New-Keynesian Regional Phillips Curve (NK-RPC) for Home PPI inflation:

$$\hat{\pi}_{H,t} = \kappa_H \hat{mc}_t^H + \beta \mathbb{E}_t[\hat{\pi}_{H,t+1}], \tag{A.148}$$

where $\hat{\pi}_{H,t}$ is the deviation of the Home PPI inflation from a steady state of zero inflation,

[8] This formulation assumes that the government has the same CES preferences over individual varieties as households.

$\kappa_H = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha}$ governs the slope of the Home NK-RPC, $\hat{mc}_t^H$ is the deviation of Home real marginal costs from the steady state marginal costs. Due to the symmetry of the problem, we can express the Foreign NK-RPC as:

$$\hat{\pi}_{F,t} = \kappa^F \hat{mc}_t^F + \beta \mathbb{E}_t[\hat{\pi}_{F,t+1}], \tag{A.149}$$

where κ_F and κ_h differ whenever the degree of price stickiness is heterogeneous between regions.

A.4.6 Government

The government consumes Home and Foreign varieties and raises revenues through lump-sum and income taxes on households. Government consumption follows an AR(1) process and its within-region composition mimics that of private consumption.[9] Per-capita total demand for Home and Foreign goods is:

$$G_t^H = (1-\rho_g)G_{ss} + \rho_g G_{t-1}^H + e_t^{g^H} + e_t^{G_h}, \tag{A.151}$$
$$G_t^F = (1-\rho_g)G_{ss} + \rho_g G_{t-1}^F + e_t^{g^F} + e_t^{G_f}, \tag{A.152}$$

where G_{ss} is the steady-state participation of government spending in output, which I assume is common between regions. The parameter ρ_g governs the persistence of fiscal shocks. Regional government spending is subject to two different idiosyncratic shocks: $e_t^{g^k}$ and $e_t^{G_k}$ for $k = H, F$. The first one is a purely region-specific shock as in Nakamura and Steinsson (2014). The process for the second one, $e_t^{G_k}$, is as follows:

$$e_t^{G_H} = \frac{s_H}{n}\varepsilon_t^G, \qquad e_t^{G_F} = \frac{1-s_H}{1-n}\varepsilon_t^G, \tag{A.153}$$

where ε_t^G is an aggregate fiscal shock that loads differently between regions. The per capita loadings are given by $\frac{s_H}{n}$ and $\frac{1-s_H}{1-n}$, respectively. The case where $s_h = n$ represents a government spending shock that hits all regions with equal intensity. The structure of this new shock mimics the Bartik-type instruments that are encountered in the local multipliers literature and results in a

[9]This implies that per-capita government spending in varieties from the Home and Foreign region are given by:

$$g_t^H(z) = G_t^H \left(\frac{p_{H,t}(z)}{P_{H,t}}\right)^{-\theta}, \qquad g_t^F(z) = G_t^F \left(\frac{p_{F,t}(z)}{P_{F,t}}\right)^{-\theta}. \tag{A.150}$$

closer map between the model and the data.[10] Lastly, total government spending G_t satisfies:

$$G_t = nG_t^H + (1-n)G_t^F. \tag{A.154}$$

The government follows a balanced budget where

$$\mathcal{P}_t G_t = T_t + \tau_t \Big(nW_{H,t}L_{H,t} + (1-n)W_{F,t}L_{F,t} \Big), \tag{A.155}$$

where T_t are the total lump sum taxes and $\mathcal{P}_t$ the price index of government purchases.

A.4.7 Monetary Policy

The monetary authority sets the nominal interest rate following a Taylor rule:

$$i_t = \rho_i i_{t-1} + (1-\rho_i)\Big(i_{ss} + \phi_\pi \Pi_t^{agg} + \phi_y \hat{y}_t^{agg} \Big) + u_{it}, \tag{A.156}$$

where $\hat{x}_t$ denotes the deviation of variable x from the zero inflation steady state. Π_t^{agg} is aggregate inflation, defined as $\Pi_t^{agg} = n\Pi_t + (1-n)\Pi_t^*$. Π_t and Π_t^* are the CPI inflation rates in the Home and Foreign region, respectively. The national output gap is defined by $\hat{y}_t^{agg} = n\hat{y}_t^H + (1-n)\hat{y}_t^F$.

A.4.8 Market Clearing

Per capita Home and Foreign output are

$$Y_t^H = \frac{1}{n}\int_0^1 y_t^H(z)dz, \qquad Y_t^F = \frac{1}{1-n}\int_0^1 y_t^F(z)dz. \tag{A.157}$$

Then, the total output is

$$Y_t = nY_t^H + (1-n)Y_t^F. \tag{A.158}$$

Clearing of the goods market requires the following.

$$nY_t^H = nC_{H,t}^H + (1-n)C_{F,t}^H + nG_t^H, \tag{A.159}$$

$$(1-n)Y_t^F = nC_{H,t}^F + (1-n)C_{F,t}^F + (1-n)G_t^F, \tag{A.160}$$

where all variables refer to amounts per capita.

[10] The reason is that in a setting where aggregate policies are not completely *differenced-out*, the size of the aggregate response becomes relevant. It therefore makes a difference whether we study a fiscal shock to a marginal region versus an aggregate and sizable increase in government spending that loads differently across regions. The aggregate responses to these two types of shocks will be different. In the extreme, a shock to an infinitesimal region would trigger an infinitesimal aggregate response, so whether this is differenced out or not will likely not matter much.

A.4.9 Calibration

Table A.9 summarizes the calibration. I stick to the calibration in Nakamura and Steinsson (2014) for most parameters. The first part of the table presents all parameters that are the same in both papers. In the second part of the table, I show parameter choices that differ from Nakamura and Steinsson (2014) or are newly introduced in this paper. The baseline calibration sets $\lambda_k = 0$ and generates heterogeneous interest rate sensitivities through differences in the IES of Ricardian households. These are set to $\sigma_H = 1.25$ and $\sigma_F = .75$. The parameter s_H, which governs the degree of Home bias in the aggregate government spending shock, ε_t^G, is set to .8. This implies that the Home region is relatively more exposed to government spending shocks.

Table A.9. Baseline Calibration.

A. First Set of Parameters		
β	Discount factor	.99
ν	Frisch Elasticity of Labor Supply	1
η	Elasticity of Substitution between Home and Foreign Goods	2
θ	Elasticity of Substitution between individual varieties	7
a	Labor Coefficient in Production Function	.63
i_{ss}	Steady-state Nominal Rate	.01
G_{ss}	Participation of Government Spending in Steady State	.2
τ	Income Tax Rate	0
ϕ_H	Home Bias in Home Region	.85
α	Degree of Price Stickiness	.75
ρ_i	Persistence of Monetary Policy Rule	.8
ϕ_π	Inflation Coefficient in Taylor Rule	1.5
ϕ_y	Output Coefficient in Taylor Rule	.1
B. Second set of Parameters		
n	Size of Home Region	.5
z	Ricardian Participation in Profits	1
λ_r	Share of *Hand-to-Mouth* in Region r	0
σ_H (σ_F)	IES of Ricardian Households in Home (Foreign) region	1.25 (.75)
s_H	Home Bias of Aggregate Government Spending Shock	.8
ρ_g	Persistence of Government Spending	.85

Appendix B

Appendix to Chapter 2

B.1 Data Cleaning

B.1.1 Bank Balance Sheet Data

Our initial dataset consisted of raw data from 64 banks operating domestically. To ensure the reliability and relevance of our analysis, we applied a series of filters which narrowed down the dataset to a final sample of 46 banks.

We exclude: i) seven banks that were not operational in 2016; ii) two banks with no lending activity according to the credit registry data in certain years of our sample; iii) four banks that were not actively operating in all years of our sample; and, iv) five banks that reported zero deposits in all years. These filters leave us with a balanced panel of banks that actively participate in the domestic financial system. Our sample accounts for 97% of total private sector deposits and 96% of total credit granted to the private sector.

B.1.2 Data Quality

We combine two sources of micro-data to construct bank credit variables. Both are collected by the Credit Registry of the Central Bank of Argentina. The first is the core information collected at the firm-bank level with monthly frequency. This data set contains outstanding end-of-month total credit levels for every bank-firm pair between 2013 and 2019. We supplement these data with an information Annex that started being recorded in 2015Q4. We source firm level data disaggregated by currency[1] from this Annex. Note that this information is not available at the firm-bank level, unlike the first dataset. Figure B.1 compares the aggregate bank credit from our two micro-data sources with the aggregate data published by the Central Bank. Our aggregate series, constructed with firm-bank data, accurately tracks the macro series. The aggregate series built from the Annex data performs relatively well in terms of tracking the slope of the macro data but misses its level. Importantly, Figure B.2 shows that total dollar debt from the Annex closely

[1]Banks separately inform the outstanding end-of-month level of debt in domestic and foreign currency. The latter is informed in US dollars.

tracks the evolution of dollar debt in the macro data. In Figure B.5 we plot both aggregate dollar debt series in first differences against each other. Most data points lie at the 45 degree line[2].

The aggregate series for Peso debt that we construct out of the Annex data has a number of caveats. Figure B.3 plots our results. First, discrepancies in the level of Peso debt account for the bulk of the difference between our total debt series and that of the macro data. Second, and most importantly, the Peso debt series from the Annex crosses the macro-data during our period of analysis.

We choose a conservative approach and avoid using the pesos debt data from the Annex. Instead, we take total firm-level debt from the core dataset, firm-level dollar debt from the Annex and combine them to get an implicit measure of firm-level pesos debt. Figure B.7 shows that, despite level differences, our implicit measure of pesos debt is well aligned with the macro data once first differenced.

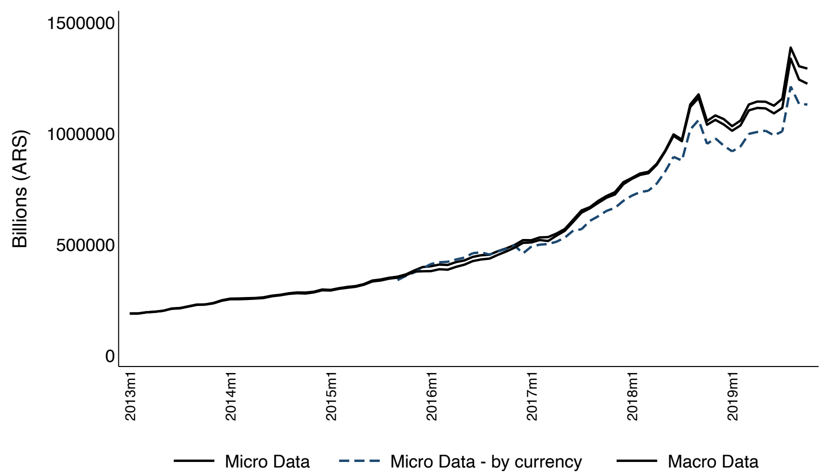


Figure B.1. Total Bank Credit

Note: Data are expressed in billions of Argentine Pesos. Macro data corresponds to the end-of-period total credit granted to private non-financial firms published by the Central Bank of Argentina. The aggregated micro data excludes firms in the public and financial sectors to maximize comparison.

[2]Since our empirical approach differences out the level of debt, this is the relevant statistic to assess the quality of our data.

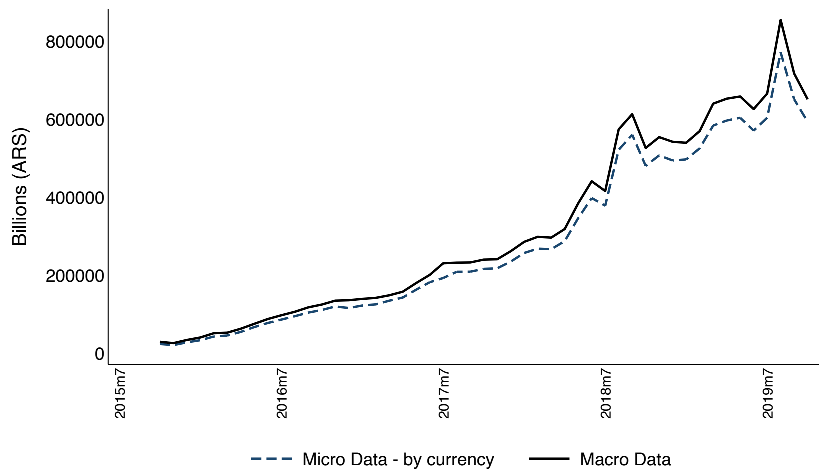


Figure B.2. Dollar Bank Credit

Note: Data are expressed in billions of Argentine Pesos. Macro data corresponds to the end-of-period total credit in foreign currency granted to private non-financial firms published by the Central Bank of Argentina. The aggregated micro data excludes firms in the public and financial sectors to maximize comparison.

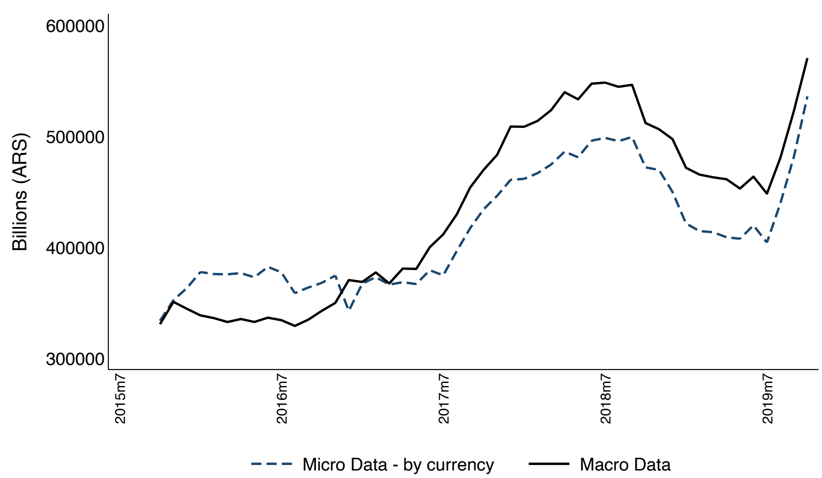


Figure B.3. Pesos Bank Credit (Raw)

Note: Data are expressed in billions of Argentine Pesos. Macro data corresponds to the end-of-period total credit in domestic currency granted to private non-financial firms published by the Central Bank of Argentina. The aggregated micro data excludes firms in the public and financial sectors to maximized comparison.

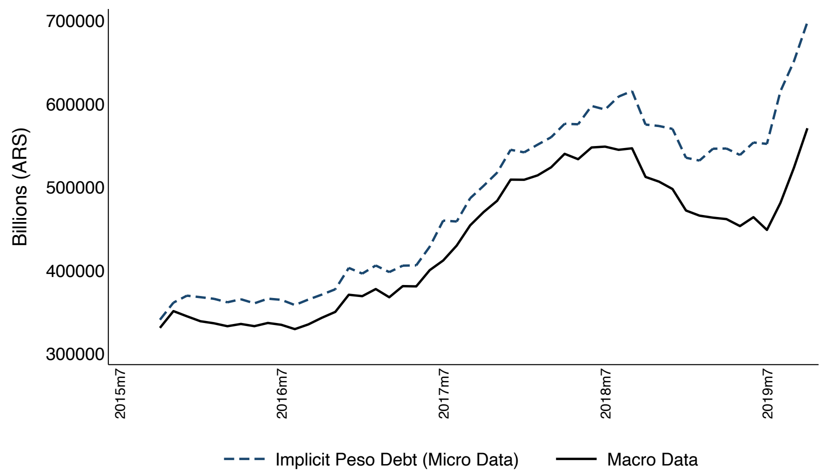


Figure B.4. Pesos Bank Credit (Implicit)

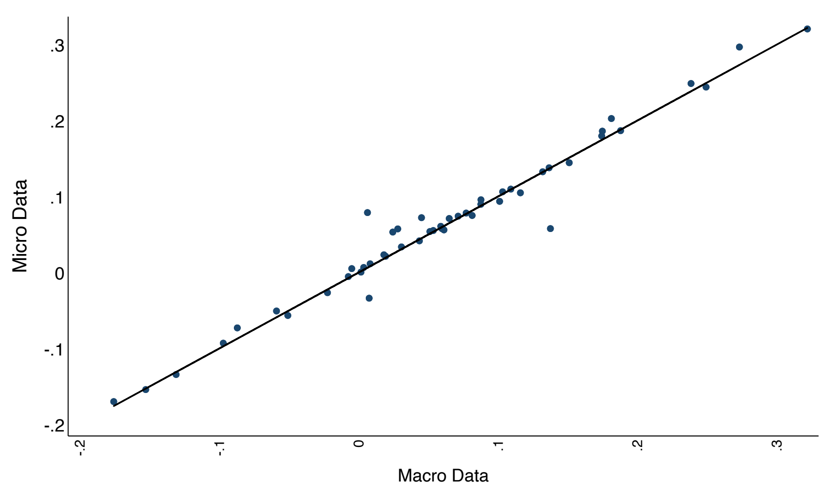


Figure B.5. Dollar Bank Credit - First Differences

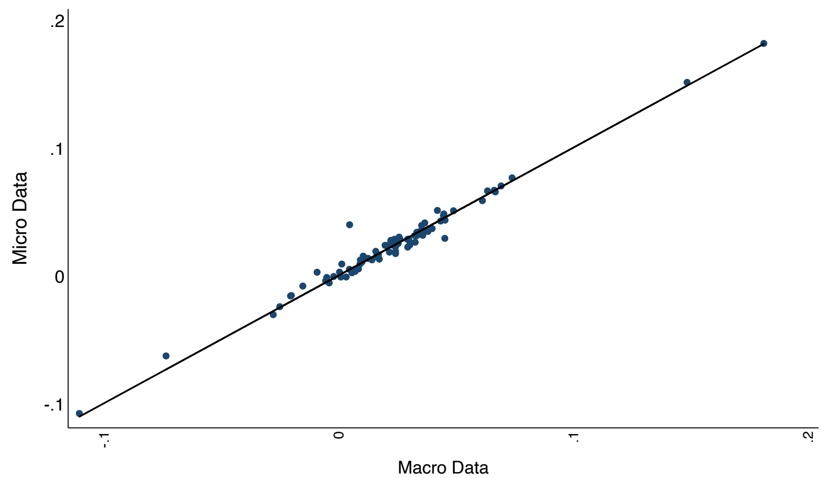


Figure B.6. Total Bank Credit - First Differences

Note: Original data was log transformed and first differenced. Micro data refers to the dollar series constructed using the Annex. Macro data corresponds to the aggregate debt series published by the Central Bank of Argentina

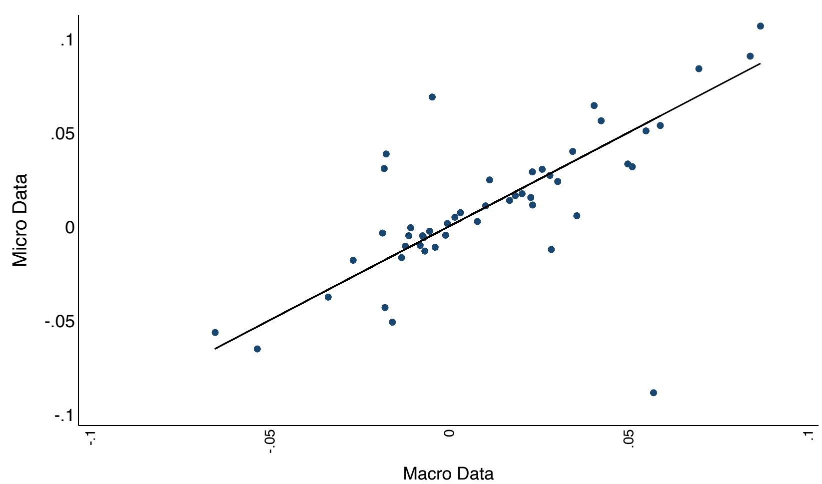


Figure B.7. Pesos (Implicit) Bank Credit - First Differences

Note: Original data was log transformed and first differenced. Micro data refers to the dollar series constructed using the Annex. Macro data corresponds to the aggregate dollar debt series published by the Central Bank of Argentina

B.1.3 Alternative Measure of Firm Exposure

Our baseline measure of firm exposure is:

$$S_i = \sum_b \omega_{ib}^{Start} S_b \tag{B.1}$$

where the weights ω_{ib}^{Start}, correspond to the participation of bank b in the total debt of firm i in the quarter before the implementation of the tax amnesty (i.e., 2016Q2). The binscatter plot below shows that our measure of firm exposure is invariant with regard to considering the average weight during the year prior to the tax amnesty. The correlation coefficient between these two alternatives is $\approx .97$.

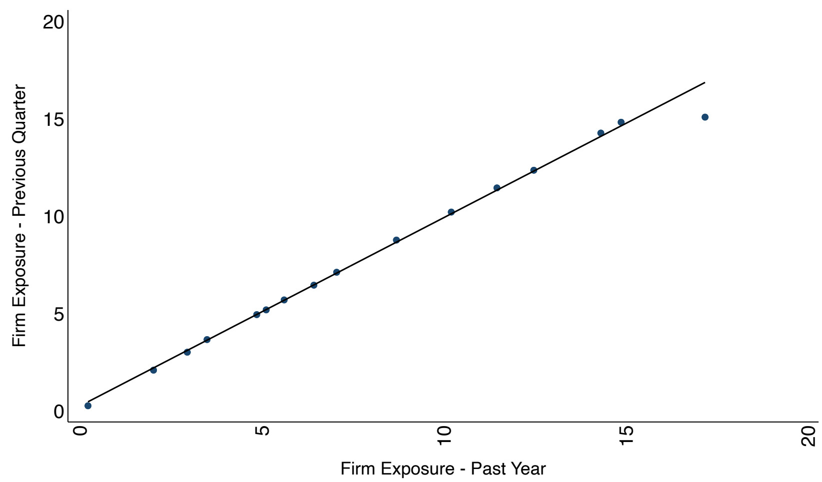


Figure B.8. Firm Exposure: Alternative Weights

Note: The *x-axis* computes firm exposure using year-to-date weights while the *y-axis* uses last quarter's weights. The correlation coefficient between these two alternative ways of measuring firm exposure is $\approx .97$.

B.2 Composition of Dollar Debt

We group companies with positive dollar debt between 2015Q4 and 2016Q4 into six mutually exclusive groups: (i) exporters with dollar debt above 1,000; (ii) exporters with dollar debt below 1,000; (iii) importers that do not export with dollar debt above 1,000; (iv) importers that do not export with dollar debt below 1,000; (v) firms that neither export nor import with dollar debt above 1,000; and (vi) firms that neither export nor import with dollar debt below 1,000. Table B.1 summarizes the characteristics of each group. Around two thousand exporting firms take on the bulk of dollar debt with an average participation of 87%. They also have higher levels of average and median dollar debt per firm. Importers who do not export represent 5.5% of the total dollar debt at the start of the study, while the rest of firms that neither export nor import take up 6.8%. The average and median dollar debt levels of both types of firms are orders of magnitude lower than those of exporters. The industrial composition differs substantially between exporters and the other two groups of firms (regardless of their dollar debt level (above/below 1K). Among exporters, 70% belong to the industrial or agricultural sector. This share drops to 30% for nonexporting importers and to only 10% among neither exporting nor importing firms.

Table B.1. Composition of Dollar Debt - at Baseline.

	Average Debt	Median Debt	Participation(%)	Manufacturing(%)	Agricultural(%)	Number of Firms
Exporters (+1k)	1243541	34892	87.4	0.6	0.1	1893
Exporters(-1k)	272	171	0.0	0.7	0.0	2076
Importers(+1k)	115554	2335	5.6	0.3	0.0	1586
Importers(-1k)	260	162	0.0	0.3	0.0	3044
Non Exp./Imp.(+1k)	54666	2202	6.8	0.1	0.1	14659
Non Exp.Imp.(-1k)	190	80	0.1	0.1	0.1	17190

Note: The table reports descriptive statistics for firms with positive dollar debt between 2015Q4 and 2016Q4. Debt values are expressed in nominal US dollars. Participation denotes each group's share of total dollar debt at baseline.

B.3 Empirical Analysis

B.3.1 Characteristics of Banks Receiving Tax Amnesty Funds

Figure B.9 shows a binscatter plot of bank exposure against the share (%) of dollar deposits before the shock.

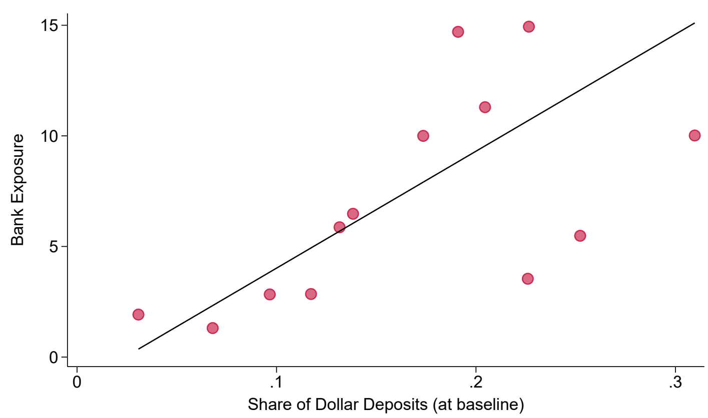


Figure B.9. Bank Exposure and Dollar Deposit Share.

Note: The share of dollar deposits in the $x-axis$ is the average share for $2016Q2$.

We also leverage the mandatory six-month holding period for tax amnesty special accounts to track the movement of these funds once they become eligible for reallocation. If households have other accounts in the bank, it is expected that they will move their dollars within the same bank. As shown in Figure 2.1B, the balances in tax amnesty accounts declined dramatically across banks after the end of this period. At the same time, the overall dollar deposits within each bank remained largely stable. This stability suggests that individuals primarily transferred funds from special accounts to their existing standard dollar accounts within the same bank. To see this, Figure B.10 presents a histogram of the percentage change in total dollar deposits between the first quarter of 2017 - when most funds were required to be held in special accounts - and the second quarter of 2017, after the holding period expired. The median change is 3.8%, and 85% of banks in the sample experienced fluctuations of less than 10%.

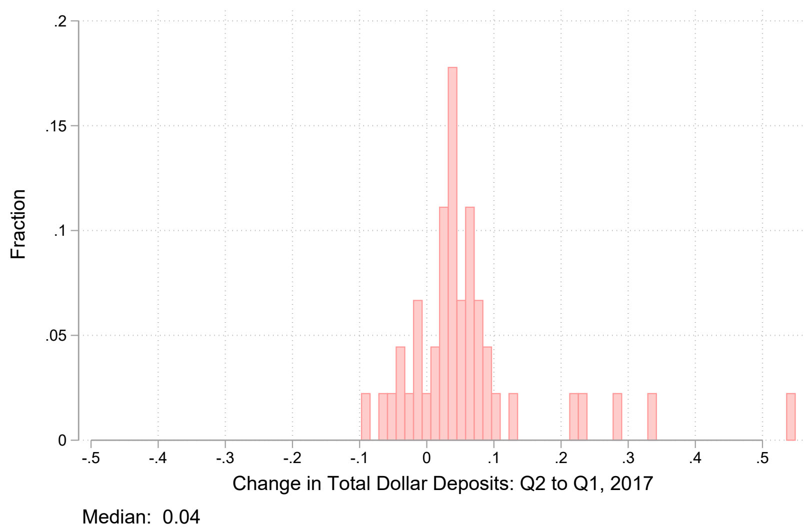


Figure B.10. Total Dollar Deposits by Bank: Changes between 2017q1-2017q2.

Note: The x-axis represents the change in total dollar deposits between 2017q2 (period in which the mandatory special accounts were not binding) and 2-17q1 (period in which the mandatory special accounts were binding).

B.4 Bank Lending Channel - Robustness Checks

This section presents robustness results for the bank lending exercise. Table B.2 shows the results when we flexibly control for bank specialization at the industry and international trade financing level.

We construct measures of bank specialization to control directly for bank-firm specific demand shocks. Specifically, we define banks' specialization along two dimensions. Industry specialization: we measure the extent to which bank b specializes in lending to a firm's industry j. We construct two measures based on 2015 (pre-shock) data: (i) the share of total lending to industry j provided by bank b, and (ii) the share of bank b's total lending dedicated to industry j. Foreign trade specialization: we similarly measure banks' specialization in lending to exporting and importing firms, constructing two analogous measures based on pre-shock data: (i) the share of total trade-related lending provided by bank b to exporters/importers, and (ii) the share of bank b's total lending directed towards exporters/importers. As a robustness check, we include the interaction between the specialization measures, the firm type (industry or trade participation) and a post-amnesty indicator in our regression specifications. This strategy allows us to directly absorb any firm-bank-specific credit demand shocks correlated with bank specialization.

Table B.3 presents additional results where we control for a battery of bank-level characteristics interacted with a post-treatment dummy. Overall, our results remain remarkably stable across specifications.

Table B.2. Bank Lending Channel - Bank Specialization Controls.

	Baseline	Industry$_A$	Exporter$_A$	Importer$_A$	Industry$_B$	Exporter$_B$	Importer$_B$
Bank Exposure × Post	1.10***	1.11***	0.99***	0.84***	1.10***	1.09***	1.11***
	(0.132)	(0.132)	(0.132)	(0.131)	(0.132)	(0.132)	(0.132)
N	1635729	1635496	1635729	1635729	1635496	1635729	1635729
Firm-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank-Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Public Bank-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank Size - Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.87	0.87	0.87	0.87	0.87	0.87	0.87

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the bank-firm level. The subscript A is used to denote the specialization measure computed as the share of credit from bank b going to firm type X, where X is either industry, exporting or importing status. The subscript B denotes the specialization measure computed as the share of credit to firm type X supplied by bank b. Point estimates measure the percent increase in L_{ibt} for every $1p.p.$ increase in bank exposure, S_b. $Post$ is equal to one between $(2016Q3, 2018Q2)$. The sample consists of all firms that have debt with two or more banks. We winsorize the top and bottom 1% of observations to limit the influence of outliers. Bank size is measured as the participation of each bank in total deposits at baseline and public bank is an indicator for a bank being publicly owned.

Table B.3. Bank Lending Channel - Bank-level Time Varying Controls.

	Baseline	Leverage	Dollar Assets	Default(%)	ROE	CB Bonds
Bank Exposure × Post	1.10***	1.00***	1.10***	1.20***	1.03***	1.34***
	(0.132)	(0.131)	(0.132)	(0.135)	(0.133)	(0.141)
N	1635729	1635729	1635729	1635729	1635729	1635729
Firm-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank-Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Public Bank-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank Size - Time FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.87	0.87	0.87	0.87	0.87	0.87

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses and clustered at the bank-firm level. Each column incorporates an additional interaction between the post indicator and a bank characteristic. Leverage measures bank leverage, dollar assets measures the participation of dollar securities in total dollar assets, default is the share of private sector loans in default, ROE measures return on equity and CB Bonds measures the participation of bonds issued by the central bank on total assets. Point estimates measure the percent increase in L_{ibt} for every $1p.p.$ increase in bank exposure, S_b. $Post$ is equal to one between $(2016Q3, 2018Q2)$. The sample consists of all firms that have debt with two or more banks. We winsorize the top and bottom 1% of observations to limit the influence of outliers. Bank size is measured as the participation of each bank in total deposits at baseline and public bank is an indicator for a bank being publicly owned.

B.5 Elasticity of Substitution Between Peso and Dollar Debt

B.5.1 Pesos Lending Rate and Bank Exposure

We tested whether bank exposure had an impact on the average peso lending rate set by individual banks. A differential response of the peso-lending rate across banks with different exposure would pose a threat to our identification assumption. Unfortunately, the only information we have on interest rates is the average lending rate for the entire portfolio of peso-denominated loans granted to the private sector. This series includes both loans to households and private nonfinancial firms. It also reflects the specific composition of credit lines in each bank portfolio that are likely to be heterogeneous. In the exercise below, we include several controls in an attempt to reduce the influence of heterogeneity on the composition of the loan portfolio on our estimates.

We run the following lag-augmented panel local projection:

$$\Delta i_{bt+h}^{Pesos} = \beta_h S_b + \Gamma_h \mathbb{X}_b + \sum_{s=1}^{2} \delta_{hs} i_{bt-s}^{Pesos} + e_{bt+h} \tag{B.2}$$

where i_{t+h}^{Pesos} is bank's b average pesos lending rate h periods after the start of the tax amnesty and S_b is bank exposure. β_h measures the differential effect of bank exposure on the pesos lending rate at horizon h. We let $h = \{-4, 8\}$, which allows us to check for pretrends. $\mathbb{X}_b$ is a vector of bank-level controls. On top of the bank controls included in the main text (i.e. bank size and bank ownership status), we add the share of commercial loans at baseline. This additional control helps us to account for differences in the composition of loan portfolios. We incorporate two lags of the dependent variable and cluster standard errors at the bank level. Figure B.11 plots the estimated coefficients. We do not find significant evidence of a differential response of pesos lending rates across banks with differential exposure to the tax amnesty shock. For reference, the average lending rate was 32.3% during this time period.

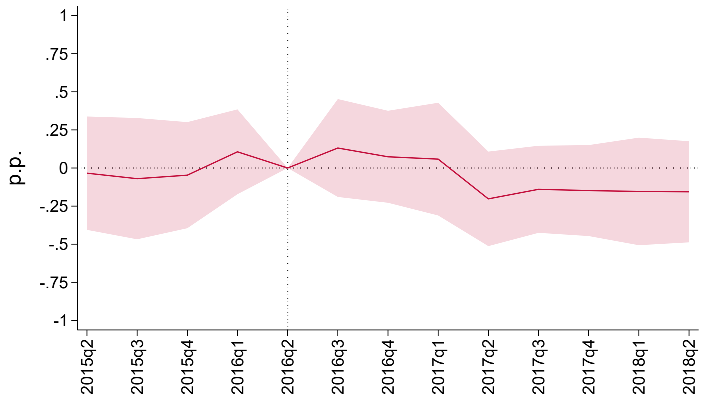


Figure B.11. Bank Exposure and Pesos Lending Rates

Note: We report 95% confidence intervals. Standard errors are clustered at the firm level.Point estimates measure the percentage point increase in the pesos lending rate for every 1p.p. increase in bank exposure, relative to baseline. The time period corresponds to 2015Q2 2018Q2 and the base period is set to 2016q2. Standard errors are clustered at the bank level.

Appendix C

Appendix to Chapter 3

C.1 Simple Model Extension

C.1.1 Hand-to-Mouth Households

We assume a fraction μ of all households are hand-to-mouth. Hand-to-mouth households are distributed across regions proportional to output across regions. They earn income a fraction μ of local income and only spend on local goods.

Market clearing for each good is,

$$Y_{i,g,t} = (1-\mu)C^{o}_{i,g,t} + \mu C^{m}_{i,g,t} \tag{C.1}$$

where $C^{o}_{i,g,t}$ is per capita consumption by the optimizing representative household and $C^{m}_{i,g,t}$ is per capita consumption by hand-to-mouth households.

Hand-to-mouth consumers have consumption shares α_{gt} for each good g, similar to optimizing households,

$$C^{m}_{i,g,t} = \alpha_{gt}\left(\frac{P_{i,g,t}}{P_{it}}\right)^{-\sigma_g} C^{m}_{it}. \tag{C.2}$$

Given optimal prices (3.15), we get

$$C^{m}_{g,t} = \alpha_g Y_t \frac{Z_{t,g}}{Z_t} \tag{C.3}$$

$$C^{o}_{g,t} = \alpha_g \left(\frac{\lambda_t}{Z_t}\right)^{-\frac{1}{\sigma_g}} Z_{t,g}. \tag{C.4}$$

Market clearing implies,

$$Y_{g,t} = (1-\mu)\alpha_g \left(\frac{\lambda_t}{Z_t}\right)^{-\frac{1}{\sigma_g}} Z_{t,g} + \alpha_g \mu Y_t \frac{Z_{t,g}}{Z_t}. \tag{C.5}$$

Total local output is,

$$\begin{aligned}
Y_{i,t} &= \sum_{g=1}^{G}(1-\mu)\alpha_{ig}e^{\hat{\alpha_{it}}}\left(\alpha_g\left(\frac{\lambda_t}{Z_t}\right)^{-\sigma_g}Z_{t,g}\right)+\mu C_{i,t}^m \\
&= e^{\hat{\alpha_{it}}}\sum_{g=1}^{G}\alpha_{ig}Y_{g,t}-e^{\hat{\alpha_{it}}}\mu Y_t\frac{Z_{it}}{Z_t}+\mu Y_{i,t} \qquad \text{(C.6)} \\
&= e^{\hat{\alpha_{it}}}\left[\frac{1}{1-\mu}\sum_{g=1}^{G}\alpha_{ig}Y_{g,t}-\frac{\mu}{1-\mu}Y_t\frac{Z_{it}}{Z_t}\right].
\end{aligned}$$

The log-linear approximation around the trend growth path is,

$$\hat{y}_{g,t} = -\frac{1}{\sigma_g}(1-\mu)\hat{\lambda}_t+\mu\hat{y}_t \tag{C.7}$$

$$\hat{y}_t = -\sum_{g=1}^{G}\left(\alpha_g\frac{1}{\sigma_g}\right)\hat{\lambda}_t. \tag{C.8}$$

Local output growth is

$$\Delta^h y_{i,t,t+h} \approx \hat{\alpha_{it}}+\frac{1}{1-\mu}\sum_{g=1}^{G}\mu_{i,g,t}\Delta^h y_{g,t+h}+\frac{\mu}{1-\mu}\Delta^h y_{t+h}. \tag{C.9}$$

The aggregate business cycle at t is unchanged,

$$\hat{y}_t+\xi\sum_{s=0}^{t-1}\hat{y}_s. \tag{C.10}$$

Therefore, the cyclical Bartik is,

$$y_{i,t,t+h}^{BC} = (1-\mu)\left(\sum_{g=1}^{G}\mu_{i,g,t}\frac{1}{\sigma_g}\right)\left(\sum_{g=1}^{G}\mu_{g,t}\frac{1}{\sigma_g}\right)^{-1}\left(\Delta^h\hat{y}_{t+h}+\xi\sum_{s=0}^{h-1}\hat{y_{t+s}}\right)+\mu\left(\Delta^h\hat{y}_{t+h}+\xi\sum_{s=0}^{h-1}\hat{y}_{t+s}\right). \tag{C.11}$$

Under conditions from Proposition 3.1, we get a first stage and reduced form coefficient equal to,

$$\gamma_1 = \frac{1}{1-\mu} \tag{C.12}$$

$$\delta_1 = \frac{1}{1-\mu}\frac{\Delta^{T+h-P}\hat{y}_{T+h}+\xi\sum_{s=0}^{T+h-P-1}\hat{y}_{P+s}}{\Delta^{T-P}\hat{y}_T+\xi\sum_{s=0}^{T-P-1}\hat{y}_{P+s}}. \tag{C.13}$$

The IV coefficient is the ratio δ_1/γ_1 and unchanged from Proposition 3.1. Intuitively, the Bartik underpredicts the recession in more exposed regions because it captures only the hand-to-mouth response in the average region. The Bartik exposure therefore gets scaled up in the first stage (and reduced form), and the common hand-to-mouth response in the average region gets absorbed by the recession fixed effect.

C.1.2 Local Prices Respond to Recession

We assume that the relative price of good g in region i relative to the industry price is set based on the relative recession in region i with elasticity κ_I,

$$\frac{P_{i,g,t}}{P_{g,t}} = \frac{\bar{W}_{ig}}{\bar{W}_g}\left(\frac{Y_{i,t}/Z_{i,t}}{Y_t/Z_t}\right)^{\kappa_I} \tag{C.14}$$

and the industry price relative to the aggregate price is based on the goods-level relative recession with elasticity $\kappa_G < \min_g \sigma_g$,

$$\frac{P_{g,t}}{P_t} = \frac{\bar{W}_g}{\bar{W}}\left(\frac{Y_{g,t}/Z_{g,t}}{Y_t/Z_t}\right)^{\kappa_G}\frac{Z_t}{Z_{g,t}} \tag{C.15}$$

where $\kappa_G = \kappa_I = 0$ (acyclical relative prices) corresponds to the baseline model.

Then local output of good g is equal to

$$Y_{i,g,t} = \alpha_{igt}\left(\frac{\bar{W}_g}{\bar{W}}\right)^{-\eta}\left(\frac{Y_{i,t}/Z_{i,t}}{Y_t/Z_t}\right)^{-\eta\kappa_I} Y_{g,t} \tag{C.16}$$

and aggregate output of good g is equal to,

$$Y_{g,t} = Z_{g,t}\left[\alpha_g\left(\frac{\bar{W}_g}{\bar{W}}\right)^{-1}\left(\frac{Y_t}{Z_t}\right)^{\kappa_G}\lambda_t^{-1}Z_t\right]^{\frac{1}{\sigma_g-\kappa_G}}. \tag{C.17}$$

Local output growth has a Bartik structure

$$\begin{aligned}\underbrace{\Delta^h y_{i,t+h}}_{\text{Local Output Growth}} &= \frac{1}{1+\eta\kappa_I}\underbrace{\Delta^h\hat{\alpha}_{i,t+h}}_{\text{Local Demand Shifter}} + \frac{1}{1+\eta\kappa_I}\underbrace{\sum_{g=1}^{G}\mu_{i,g,t}\Delta^h y_{g,t+h}}_{\text{Bartik Predicted Growth, } y^B_{i,t,t+h}} \\ &+ \frac{\eta\kappa_I}{1+\eta\kappa_I}\underbrace{\Delta^h y_{t+h}}_{\text{Aggregate Growth}}.\end{aligned} \tag{C.18}$$

Define $\tilde{\sigma}_g \equiv \sigma_g - \kappa_G > 0$. The cyclical component of industry growth is,

$$\hat{y}_{g,t} + \xi \sum_{s=0}^{t-1} \hat{y}_{g,s} = \frac{\kappa_G}{\tilde{\sigma}_g}\left(\hat{y}_t + \xi \sum_{s=0}^{t-1} \hat{y}_s\right) - \frac{1}{\tilde{\sigma}_g}\left(\hat{\lambda}_t + \xi \sum_{s=0}^{t-1} \hat{\lambda}_s\right). \tag{C.19}$$

The cyclical component of output is,

$$\hat{y}_t + \xi \sum_{s=0}^{t-1} \hat{y}_s = -\left(\frac{\sum_{g=1}^{G} \mu_{gt} \frac{1}{\tilde{\sigma}_g}}{1 - \kappa_G \sum_{g=1}^{G} \mu_{gt} \frac{1}{\tilde{\sigma}_g}}\right)\left(\hat{\lambda}_t + \xi \sum_{s=0}^{t-1} \hat{\lambda}_s\right) \tag{C.20}$$

which implies,

$$\hat{y}_{g,t} + \xi \sum_{s=0}^{t-1} \hat{y}_{g,s} = \frac{1}{\tilde{\sigma}_g}\left[\kappa_G + \frac{1 - \kappa_G \sum_{k=1}^{G} \mu_{kt} \frac{1}{\tilde{\sigma}_k}}{\sum_{k=1}^{G} \mu_{kt} \frac{1}{\tilde{\sigma}_k}}\right]\left(\hat{y}_t + \xi \sum_{s=0}^{t-1} \hat{y}_s\right). \tag{C.21}$$

The cyclical component of Bartik predicted growth is,

$$y_{i,t,t+h}^{BC} = \underbrace{\left(\frac{\sum_{g=1}^{G} \mu_{i,g,t} \frac{1}{\tilde{\sigma}_g}}{\sum_{k=1}^{G} \mu_{kt} \frac{1}{\tilde{\sigma}_k}}\right)}_{\text{Relative Regional Business Cycle Exposure}} \times \underbrace{\left(\Delta^h \hat{y}_{t+h} + \xi \sum_{s=0}^{h-1} \hat{y}_{t+s}\right)}_{\text{Aggregate Business Cycle}}. \tag{C.22}$$

Under conditions from Proposition 3.1, we get a first and reduced form coefficient equal to,

$$\gamma_1 = \frac{1}{1 + \eta \kappa_I} \tag{C.23}$$

$$\delta_1 = \frac{1}{1 + \eta \kappa_I} \frac{\Delta^{T+h-P} \hat{y}_{T+h} + \xi \sum_{s=0}^{T+h-P-1} \hat{y}_{P+s}}{\Delta^{T-P} \hat{y}_T + \xi \sum_{s=0}^{T-P-1} \hat{y}_{P+s}}. \tag{C.24}$$

The IV coefficient is the ratio δ_1/γ_1 and unchanged from Proposition 3.1.

The Bartik now overpredicts the recession in more exposed regions because it captures only the price response in the average region. The Bartik exposure therefore gets scaled down in the first stage (and reduced form). Good-specific price responses do not pose an issue as these only affect how the aggregate shock affects the Bartik, but not how the Bartik differentially affects regions.

C.1.3 Endogenous Monetary Policy Response

Suppose the central bank follows a Taylor rule in which the real interest rate responds to the cyclical component of output,

$$\hat{r}_t = \phi \hat{y}_t + \hat{v}_t, \tag{C.25}$$

where $\hat{\nu}_t$ is a monetary policy shock.

Then the only difference from the baseline model is the solution for $\hat{\lambda}_t$,

$$
\begin{aligned}
\hat{\lambda}_t = \frac{1}{1+(\phi-\xi)\sum_{g=1}^{G}\frac{\alpha_g}{\sigma_g}} \sum_{s=0}^{\infty}\left(\frac{1}{1+(\phi-\xi)\sum_{g=1}^{G}\frac{\alpha_g}{\sigma_g}}\right)^{s}(\hat{\beta}_{t+s}+\hat{\nu}_{t+s}) \\
+\lim_{T\to\infty}\left(\frac{1}{1+(\phi-\xi)\sum_{g=1}^{G}\frac{\alpha_g}{\sigma_g}}\right)^{T}\mathbb{E}_t\lambda_T.
\end{aligned}
\tag{C.26}
$$

Since no other equation changes, Proposition 3.1 goes through unchanged.

C.2 Trend Mismeasurement

C.2.1 Trend Mismeasurement in the Cross-Region Approach

The IV strategy discussed in Section 3.2 requires an estimate of the Bartik predicted recession depth which we denote by $\tilde{y}^{BC}_{i,P,T}$:

$$
\tilde{y}^{BC}_{i,P,T} = \sum_{g}\mu_{i,g,P}\Delta^{P-T}\tilde{\tilde{y}}_{g,P}. \tag{C.27}
$$

We construct the estimated cyclical component of sectoral output as

$$
\tilde{\tilde{y}}_{g,t} = y_{g,t} - \tilde{\bar{y}}_{g,t}, \tag{C.28}
$$

where $\tilde{\bar{y}}_{g,t}$ is an estimate for the trend-component of sectoral output. We model the true sectoral trend as

$$
\bar{y}_{g,t} = \gamma_g t + u_{g,t} + \eta_g u_t + u_t, \tag{C.29}
$$

where $\gamma_g t$ captures a deterministic trend, $u_{g,t}$ is a sector-specific stochastic component, and u_t is an aggregate trend shock. We assume that both processes are mean-zero, $E(u_{g,t}) = E(u_t) = 0$. The parameter $\eta_g \sim (0, \sigma^2_\eta)$ governs the differential exposure of sector g to the aggregate trend shock.

Throughout, we assume that sector-specific shocks are (i) uncorrelated across goods, (ii) uncorrelated with the aggregate shock, and (iii) orthogonal to sectoral trend parameters and specialization patterns, $\{\gamma_g, \eta_g\}_g$ and $\{\mu_{i,g}\}_i$.

Assumption C.1.

$$
\begin{aligned}
&u_{g,t} \perp u_{g',s} \quad \forall\, g, g' \quad and \quad \forall\, s \\
&u_{g,t} \perp u_s \quad \forall g \in G \quad and \quad s \in T \\
&u_{g,t} \perp \{\{\gamma_g, \eta_g\}_g, \{\mu_{i,g}\}_i\} \\
&E(u_{g,t}) = E(u_t) = 0 \quad \forall\, g, s.
\end{aligned}
$$

Trend mismeasurement $\bar{e}_{g,t}$ after fitting a linear trend is

$$\bar{e}_{g,t} = (\hat{\gamma}_g - \gamma_g)t + u_{g,t} + \eta_g u_t + u_t \tag{C.30}$$

where $\hat{\gamma}_g$ is the OLS estimate for the deterministic trend coefficient. As $T \longrightarrow \infty$, $\hat{\gamma}_g - \gamma_g \longrightarrow 0$ and trend mismeasurement collapses to

$$\bar{e}_{g,t} \longrightarrow u_{g,t} + \eta_g u_t + u_t. \tag{C.31}$$

Then, the estimated Bartik instrument $\tilde{y}^{BC}_{i,P,T}$ can be expressed as

$$\tilde{y}^{BC}_{i,P,T} = \sum_g \mu_{i,g,P}[\Delta^{P-T}\hat{y}_{g,P} - \Delta^{P-T}\bar{e}_{g,P}], \tag{C.32}$$

and the IV estimator of the *relative recession impulse response function* as

$$\hat{\beta}^{IV}_h = \frac{Cov_i\Big(\Delta^{T+h-P} y_{i,T+h}, \tilde{y}^{BC}_{i,P,T}\Big)}{Cov_i\Big(\Delta^{P-T} y_{i,P}, \tilde{y}^{BC}_{i,P,T}\Big)}. \tag{C.33}$$

Under the assumptions stated in Proposition 3.1 and further assuming that trend mismeasurement $\Delta^{P-T}\bar{e}_{g,P}$ over the recession is uncorrelated with the cyclical component of sectoral output, it follows that:

$$\hat{\beta}^{IV}_h = \frac{Cov_i\Big(\Delta^{T+h-P}\hat{y}_{i,T+h}, y^{BC}_{i,P,T}\Big) - Cov_i\Big(\Delta^{T+h-P}\bar{y}_{i,T+h}, \sum_g \mu_{i,g,t}\Delta^{P-T}\bar{e}_{g,P}\Big)}{Cov_i\Big(\Delta^{P-T}\hat{y}_{i,P}, y^{BC}_{i,P,T}\Big) - Cov_i\Big(\Delta^{P-T}\bar{y}_{i,P}, \sum_g \mu_{i,g,t}\Delta^{P-T}\bar{e}_{g,P}\Big)}. \tag{C.34}$$

Assumption C.2.

$$Cov_i\Big(\sum_g \mu_{i,g,t}\Delta^{T+s-P}\hat{y}_{g,T+s}, \sum_g \mu_{i,g,t}\Delta^{P-T}\bar{e}_{g,P}) = 0 \quad \forall \quad s.$$

Then, it follows that $\hat{\beta}_h^{IV} \to \beta_h$ if, and only if

$$Cov_i\Big(\Delta^{T+s-P}\bar{y}_{i,T+s}, \sum_g \mu_{i,g,t}\Delta^{T-P}\bar{e}_{g,T}\Big) \to 0 \quad \text{for } s = 0, h, \tag{C.35}$$

where the last expression uses $\Delta^{P-T}\bar{e}_{g,P} = -\Delta^{T-P}\bar{e}_{g,T}$. This is the identifying assumption for trend mismeasurement to wash-out in the regional design. The covariance in (C.35) can be re-written as:

$$Cov_g\Big(\Delta^D\bar{e}_{g,T}, \sum_i \mu_{i,g}\Delta^{D+s}\bar{y}_{i,T+s}\Big) \to 0, \tag{C.36}$$

where $D = T - P$ denotes recession duration. For $T \to \infty$, replacing for the definitions of $\bar{y}_{i,T+s}$ and $\bar{e}_{g,t}$ yields:

$$\begin{aligned} &Cov_g\Big(\Delta^D u_{g,T} + \eta_g\Delta^D u_T + \Delta^D u_T, \\ &\qquad \sum_{g'}\Theta_{g,g'}\Big\{\gamma_{g'} \times (D+s) + \Delta^{D+s}u_{g',T+s} + \eta_{g'}\Delta^{D+s}u_{T+h} + \Delta^{D+s}u_{T+h}\Big\}\Big) \to 0, \end{aligned}$$

where $\Theta_{g,g'} = \sum_i \mu_{i,g}\mu_{i,g'}$. This covariance can be further broken up into the following six terms:

$$C_1 = (D+s)\sum_{g'}\gamma_{g'}E_g(\Theta_{g,g'}\Delta^D u_{g,T}) \tag{C.37}$$

$$C_2 = \Delta^D u_T(D+s) \times \sum_{g'}\gamma_{g'}\Big[E_g(\Theta_{g,g'}\eta_g) - \bar{\eta}E_g(\Theta_{g,g'})\Big] \tag{C.38}$$

$$C_3 = E_g(\Delta^D u_{g,T}\Delta^{D+s}u_{g,T+s}) \times E_g\Big(\sum_i \mu_{i,g}^2\Big) \tag{C.39}$$

$$C_4 = \Delta^{D+s}u_{T+s} \times E_g\Big(\Delta^D u_{g,T}\sum_{g'}\Theta_{g,g'}\eta_{g'}\Big) \tag{C.40}$$

$$C_5 = \Delta^D u_T \times \sum_{g'}\Delta^{D+s}u_{g',T+s}\Big[E_g\Big(\Theta_{g,g'}\eta_g\Big) - \bar{\eta}E_g\Big(\Theta_{g,g'}\Big)\Big] \tag{C.41}$$

$$C_6 = \Delta^D u_T\Delta^{D+s}u_{T+s} \times \underbrace{\sum_{g'}\eta_{g'}\Big[E_g\Big(\Theta_{g,g'}\eta_g\Big) - \bar{\eta}E_g\Big(\Theta_{g,g'}\Big)\Big]}_{Var_i(\eta_i)}, \tag{C.42}$$

where $\bar{\eta} = \frac{1}{G}\sum_g \eta_g$ and $\eta_i = \sum_g \mu_{i,g,t}\eta_g$. We omit the covariance terms that involve the average effect of the aggregate trend shock as those are differenced-out. First, C_1 and C_4 converge to zero by Assumption C.1—trend shocks have mean zero and are orthogonal to sectoral parameters. Second, for C_3 to converge to zero we impose the following assumption on regional output composition.

Assumption C.3.

$$E_g\Big(\sum_i \mu_{i,g,t}^2\Big) \to 0 \quad as \quad G \to \infty.$$

This asymptotic condition guarantees that sectoral-level trend shocks *wash-out* in the cross-section and can be thought of as a bound on regional specialization. In other words, we require that regions are sufficiently diversified across sectors g. This is a necessary condition.

Lastly, for C_2, C_5 and C_6 to converge to zero it suffices to establish conditions under which the term in brackets—$E_g\Big(\Theta_{g,g'}\eta_g\Big) - \bar{\eta}E_g\Big(\Theta_{g,g'}\Big)$—cancels out, although there exist weaker conditions specific to each covariance term that could make orthogonality hold. Below we establish a condition on sectoral loadings η_g that guarantees that these covariances collapse to zero:

Assumption C.4. *The sectoral-level loadings η_g are independently and identically distributed across goods with mean $\bar{\eta}$ and variance σ_η^2:*

$$\eta_g \overset{iid}{\sim} (\bar{\eta}, \sigma_\eta^2).$$

A corollary of Assumption C.4 is that regions are in expectation homogeneously exposed to the aggregate trend shock—$E(\eta_i) = \sum_g \mu_{i,g,t}E(\eta_g) = \bar{\eta}$.

C.2.2 Trend Measurement in a Time-Series Approach

We observe J recession episodes indexed by j, each with peak P_j and trough T_j. Consider a local projection for each horizon h:

$$\underbrace{y_{T_j+h} - y_{P_j}}_{\Delta^{T_j+h-P_j} y_{T_j+h}} \ = \ \alpha \ + \ \beta_h^{ts}(\hat{y}_{P_j} - \hat{y}_{T_j}) \ + \ \varepsilon_j,$$

where $\hat{y}_{P_j} - \hat{y}_{T_j}$ measures the national recession depth relative to trend. $\{\beta_h^{ts}\}_h$ is the aggregate recession impulse response function and measures the behavior of aggregate output to a 1% deeper recession. We assume that each recession episode j is driven by an i.i.d shock $\beta_{0,j} \sim (0, \sigma_\beta^2)$. This generates the cross-episode variation in recession depth required to identify $\{\beta_h^{ts}\}_h$ using a time-series approach.

Constructing recession depth relative to trend requires an estimate of the aggregate trend. The estimated trend component of aggregate output is

$$\bar{y}_t^e = \bar{y}_t + \bar{e}_t, \tag{C.43}$$

where $\bar{y}_t$ is the true trend component, $\bar{e}_t$ is measurement error and the superscript e is used to denote *estimated* components of output. The estimated cyclical component is then given by

$$\begin{aligned} \hat{y}^e_t &= y_t - \bar{y}^e_t \\ &= \hat{y}_t - \bar{e}_t, \end{aligned} \tag{C.44}$$

where $\hat{y}_t$ is the true cyclical component. The estimated h-period difference in the cyclical component is

$$\Delta^h \hat{y}^e_{t+h} = \hat{y}^e_{t+h} - \hat{y}^e_t = \Delta^h \hat{y}_{t+h} - \Delta^h \bar{e}_{t+h}. \tag{C.45}$$

The OLS estimator of the aggregate recession impulse response function converges to

$$\begin{aligned} \hat{\beta}^{ts}_h &= \frac{Cov_j(y_{T_j+h} - y_{P_j}, \hat{y}^e_{P_j} - \hat{y}^e_{T_j})}{Var_j(\hat{y}^e_{P_j} - \hat{y}^e_{T_j})} \\ &= -\frac{Cov_j(\Delta^{T_j+h-P_j}\hat{y}_{T_j+h} + \Delta^{T_j+h-P_j}\bar{y}_{T_j+h}, \Delta^{T_j-P_j}\hat{y}_{T_j} - \Delta^{T_j-P_j}\bar{e}_{T_j})}{Var_j(\Delta^{T_j-P_j}\hat{y}^e_{T_j})}. \end{aligned} \tag{C.46}$$

Further assuming that the trend and cyclical component of aggregate output are orthogonal to each other we get:

$$\hat{\beta}^{ts}_h = -\frac{Cov_j(\Delta^{T_j+h-P_j}\hat{y}_{T_j+h}, \Delta^{T_j-P_j}\hat{y}_{T_j})}{Var_j(\Delta^{T_j-P_j}\hat{y}^e_{T_j})} + \frac{Cov_j\left(\Delta^{T_j+h-P_j} y_{T_j+h}, \Delta^{T_j-P_j}\bar{e}_{T_j}\right)}{Var_j(\Delta^{T_j-P_j}\hat{y}^e_{T_j})}. \tag{C.47}$$

Proposition C.1.

$$\beta^{ts}_h \to -\frac{Cov_j(\Delta^{T_j+h-P_j}\hat{y}_{T_j+h}, \Delta^{T_j-P_j}\hat{y}_{T_j})}{Var_j(\Delta^{T_j-P_j}\hat{y}^e_{T_j})} = \beta^{ts}_h \times \kappa, \tag{C.48}$$

where $\kappa = \frac{Var_j(\Delta^{P_j-T_J}\hat{y}_{P_j})}{Var_j(\Delta^{P_j-T_J}\hat{y}^e_{P_j}))}$ *measures the degree of attenuation bias.*

The assumption that $Cov_j(\Delta^{T_j+h-P_j} y_{T_j+h}, \Delta^{T_j-P_j}\bar{e}_{T_j}\Big) = 0$ is strong. Even if the estimated trend $\bar{y}^e_t$ is orthogonal to the cycle, the covariance remains non-zero whenever the aggregate trend has a stochastic component. This is because trend growth during the recession window $[P_j, T_j]$ and trend growth over the recovery horizon $[P_j, T_j + h]$ share a common component, generating a non-zero covariance across episodes j. We show this formally next when the trend has a deterministic and stochastic component.

Let the aggregate trend $\bar{y}_t$ evolve according to

$$\bar{y}_t = \gamma t + \upsilon_t, \tag{C.49}$$

where $\upsilon_t = \rho_\upsilon \upsilon_{t-1} + \eta_t$ and $\eta_t \sim N(0, \sigma_\eta^2)$. The OLS estimator converges to

$$\bar{y}_t^e \to \hat{\gamma} t. \tag{C.50}$$

Measurement error is given by $\bar{e}_t = u_t + \nu_t$ where $\nu_t = (\hat{\gamma} - \gamma)t$ is finite-sample estimation noise (i.e., $\nu_t \to 0$ as $T \to \infty$). Under $\bar{y}_t^e \perp \{\hat{y}_s\}_s$, the covariance between output growth over the recovery and measurement error reduces to

$$\begin{aligned} Cov_j(\Delta^{D_j+h}\bar{y}_{T_j+h}, \Delta^{D_j}\bar{e}_{T_j}\Big) &= Cov_j(\gamma(D_j+h) + \Delta^{D_j+h}\upsilon_{T_j+h}, (\hat{\gamma}-\gamma)(D_j) + \Delta^{D_j}\upsilon_{T_j}\Big) \\ &= Cov_j(\gamma(D_j+h) + \Delta^{D_j+h}\upsilon_{T_j+h}, (\hat{\gamma}-\gamma)D_j + \Delta^{D_j}\upsilon_{T_j}\Big) \\ &= \gamma(\hat{\gamma}-\gamma)Var_j(D_j) + Cov_j(\Delta^{D_j+h}\upsilon_{T_j+h}, \Delta^{D_j}\upsilon_{T_j}\Big), \end{aligned} \tag{C.51}$$

where $D_j = T_j - P_j$ is recession duration. The first covariance term collapses to zero as $T \to \infty$. The second term is the main driver of identification failure. Further noting that

$$\Delta^{D_j+h}\upsilon_{T_j+h} = (\rho_\upsilon^{D_j+h} - 1)\upsilon_{P_j} + \sum_{s=1}^{D_j+h} \rho_\upsilon^{D_j+h-s} \eta_{P_j+s}, \tag{C.52}$$

$$\Delta^{D_j}\upsilon_{T_j} = (\rho_\upsilon^{D_j} - 1)\upsilon_{P_j} + \sum_{s=1}^{D_j} \rho_\upsilon^{D_j-s} \eta_{P_j+s}. \tag{C.53}$$

It follows that $Cov_j(\Delta^{D_j+h}\upsilon_{T_j+h}, \Delta^{D_j}\upsilon_{T_j}\Big) \neq 0$ whenever $\sigma_\eta^2 \neq 0$.

C.3 Figures

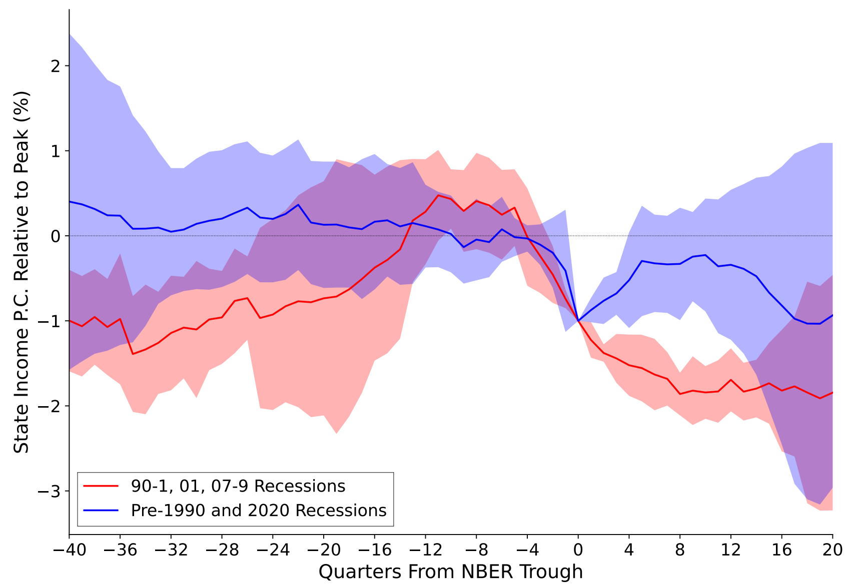
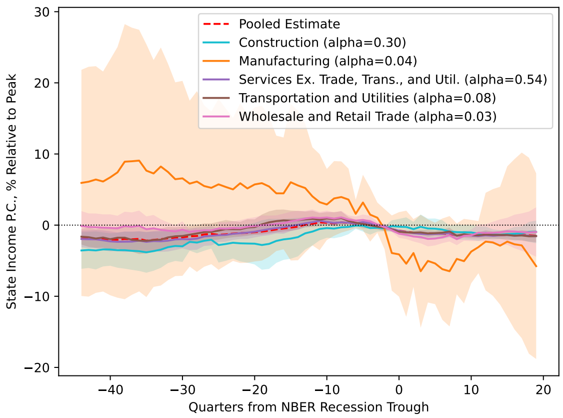
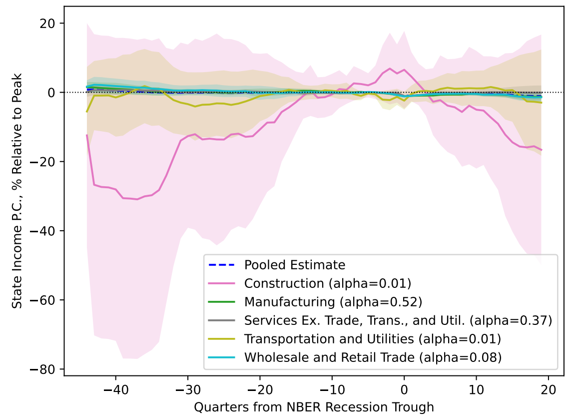


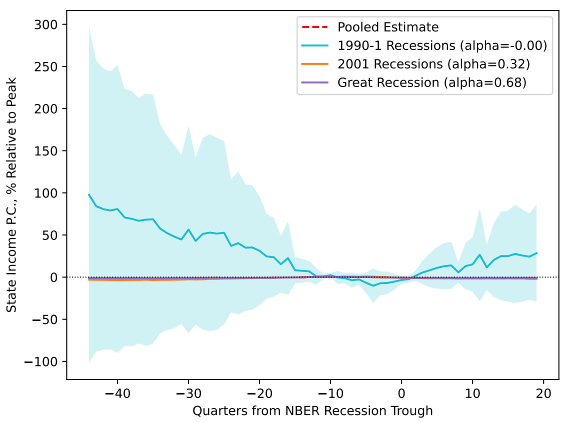
Figure C.1. IRF of State Income by Recession Group Double-Clustering by State and Recession.



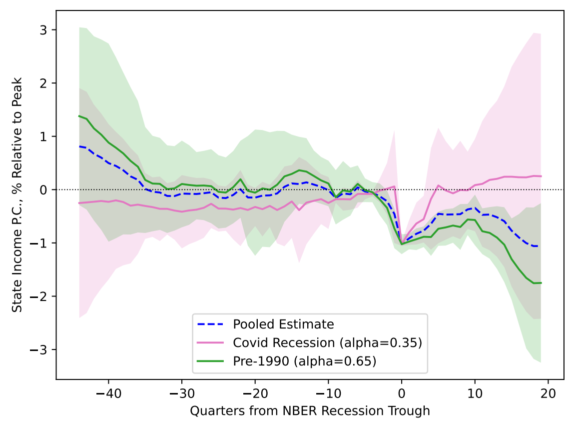
(A) Industry: 1990-1, 2001, 2007-9.



(B) Industry: Other Recessions.



(C) Recessions: 1990-1, 2001, 2007-9.



(D) Recessions: Other Recessions.

Figure C.2. IRFs based on Major Industry and Recession Variation.

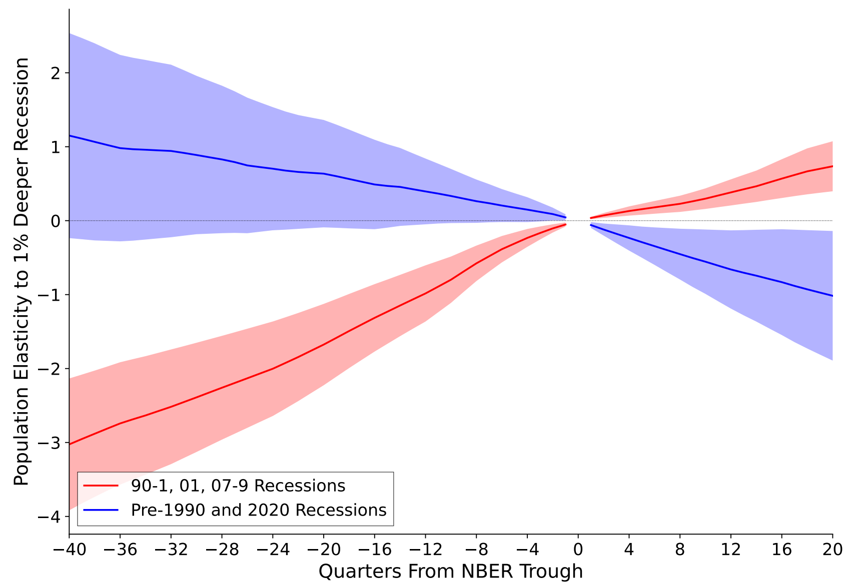


Figure C.3. Recession Impulse Response Function for Population.

C.4 Information Structure

This appendix details how we model boom-bust or liquidationist recessions using a fake news shock to household's taste for durables ψ_t.

Notation

Let $J^{i,o}$ be the full-information Jacobian for an input i and output o. Each column s of this matrix tracks the response of output o to a one-time innovation to input i in period s. Concretely, we have

$$J^{o,i} = \begin{pmatrix} \frac{\partial o_0}{\partial i_0} & \cdots & \frac{\partial o_{T-1}}{\partial i_0} \\ & \ddots & \\ \frac{\partial o_{T-1}}{\partial i_0} & \cdots & \frac{\partial o_{T-1}}{\partial i_{T-1}} \end{pmatrix} \qquad J^{o,i}_{t,s} = \frac{\partial o_t}{\partial i_s} \tag{C.54}$$

Throughout, we fix $i = \psi$ and $o = y_t$ so we omit superscripts. We use the notation $J_{[a \bullet b, c \bullet d]}$ to denote the submatrix of J from row a to b and column c to d, and is zero otherwise. We label the first column and row as 0, so the final column and row are labeled $T-1$. We also define 0_m as a row of m zeros and $0_{m,n}$ as an $m \times n$ matrix of zeros. The taste shock is a $T \times 1$ vector $\psi_{0,T-1} = (\psi_t)_{t=0}^{T-1}$. We denote the subsequence from t to $t+k$ by $\psi_{t,t+k}$. Under full information the output sequence, $y_{0,T-1}$, in response to the shock path $\psi_{0,T-1}$ is:

$$y_{0,T-1} = J\psi_{0,T-1} \tag{C.55}$$

Importantly, we can construct impulse responses that incorporate different information structures by combining different columns of the full information Jacobian J. Below we detail our approach.

Fake News Shocks

We assume the following two-wave information structure to model the fake news shock. In period $t = 0$, news arrive about an increase in ψ_t starting at time $s > 0$. This is the first wave of information. We denote the shock sequence associated to this wave by $\psi^0_{0,T-1}$:

$$\psi^0_{0,T-1} = \begin{pmatrix} 0_{0,s-1} \\ \bar{\psi}_{s,T-1} \end{pmatrix} \tag{C.56}$$

where $\bar{\psi}_k > 0$ for some $k \geq s$. Starting at $t = 0$, agents learn of the sequence $\psi^0_{0,T-1}$ at rate $1-\theta$, where θ governs the degree of information stickiness. Households that update their information set expect a higher return from durable consumption beginning in period s. In anticipation of this, they build-up their durable stock which boosts output and the investment share. Because households update their information set infrequently this results in a gradual build-up of the aggregate durable stock. At time $z < s$ a new wave of information starts. In each period $t \geq z$, a share $1-\theta$ of the agents that expect the shock sequence to be $\bar{\psi}_{t,T-1}$ learn that the sequence will instead be $\psi_{t,T-1} = \mathbf{0}$. This is equivalent to learning of a shock with path given by $\tilde{\psi}_{t,T-1} = -\psi^0_{t,T-1}$. Households that receive this new piece of information, find themselves with an excess stock of durable goods that they wish to liquidate. Initially, this results in a deceleration of the accumulation process of durable goods at the aggregate level. Accumulation decelerates but does not come to an immediate halt. This is because the stock of households that still expect the shock sequence $\psi^0_{0,T-1}$ is large relative to the stock that learns of $\tilde{\psi}_{0,T-1}$. As the second wave of information disseminates, the liquidation of excess stocks generates a bust in output, a sustained decrease in the investment share and a slow recovery of the economy relative to the preceding boom period. Next, we provide further details on how we implement and put together each of these information waves.

Implementation of the first wave

We can construct the outcome of the first information wave as the following series of learning steps:

$$y_{0,T-1} = (1-\theta)J\psi^0_{0,T-1} \qquad \text{[learn at 0]} \tag{C.57}$$

$$+\theta(1-\theta)\begin{pmatrix} 0_{1,1} & 0_{1,T-1} \\ 0_{T-1,1} & J_{[0\bullet T-2,0\bullet T-2]} \end{pmatrix}\psi^0_{0,T-1} \qquad \text{[learn at 1]} \tag{C.58}$$

$$+\theta^2(1-\theta)\begin{pmatrix} 0_{2,2} & 0_{2,T-2} \\ 0_{T-2,2} & J_{[0\bullet T-2,0\bullet T-2]} \end{pmatrix}\psi^0_{0,T-1} \qquad \text{[learn at 2]} \tag{C.59}$$

$$\vdots \tag{C.60}$$

$$+\theta^{T-1}(1-\theta)\begin{pmatrix} 0_{T-1,T-1} & 0_{T-1,1} \\ 0_{1,T-1} & J_{[0\bullet 0,0\bullet 0]} \end{pmatrix}\psi^0_{0,T-1} \qquad \text{[learn at } T-1\text{]} \tag{C.61}$$

The summation representation is given by:

$$y_{0,T-1} = \left[\sum_{k=0}^{T-1} \theta^k (1-\theta) \begin{pmatrix} 0_{k,k} & 0_{k,T-k} \\ 0_{T-k,k} & J_{[0\bullet T-1-k,0\bullet T-1-k]} \end{pmatrix}\right] \psi^0_{0,T-1} \tag{C.62}$$

where each term k in the summation captures the response of households that learn about the shock sequence $\psi^0_{0,T-1}$ in period $t = k$. We can define the sticky-information Jacobian S such that the outcome effect of the first wave is then

$$y_{0,T-1} = S\psi^0_{0,T-1} \tag{C.63}$$

S can be constructed recursively using

$$S_{[k\bullet T-1,k\bullet T-1]} = S_{[k\bullet T-1,k\bullet T-1]} + \theta^k(1-\theta) J_{[0\bullet T-1-k,0\bullet T-1-k]}, \qquad k = 0, ..., T-1 \tag{C.64}$$

$$\tag{C.65}$$

Implementation of the second wave

The second wave of information, which starts at $z < s$, gradually updates the information set of households that expect the shock sequence $\psi^0_{0,T-1}$ by feeding in the sequence $\tilde{\psi}_{0,T-1} = -\psi^0_{0,T-1}$. We refer to the stock of households that believe the shock sequence to be $\psi^0_{0,T-1}$ as believers. In each period $t \geq z$, a portion of believers learns that the sequence will be $\psi_{0,T-1} = 0$ and the remaining believers only learn of the current realization $\tilde{\psi}_t$. This means the latter know that the taste for durables did not change at t, $\psi_t = 0$, but still expect ψ^0_{t+k} for $k > 0$.

For implementation, we need to keep track of the stock of households that expect $\psi^0_{0,T-1}$ at each point in time. Every period the flow of new believers is given by:

$$\text{new believers}_t = (1-\theta)\theta^t, \qquad t = 0, ..., T-1 \tag{C.66}$$

These are the households that continue to learn about the first wave of information. Next, we assume that belief updating from the second wave of information happens immediately after the first wave and, as mentioned before, it can only affect the current stock of believers. Current believers are the sum of past believers and new believers. Starting from $t \geq z$, each period a fraction $(1-\theta)$ of them gets converted to non-believers (i.e. learns about the second wave):

$$\text{new non-believers}_t = \begin{cases} 0, & \text{if} \quad t < s \\ (1-\theta)(\text{believers}_{t-1} + \text{new believers}_t), & \text{if} \quad t \geq z \end{cases} \tag{C.67}$$

This means that the stock of believers is simply past believers plus the net change in believers:

$$\text{believers}_t = \text{believers}_{t-1} + \text{new believers}_t - \text{new non-believers}_t \tag{C.68}$$

We can compute the sticky information Jacobian associated to the second wave as follows:

$$
\begin{aligned}
\tilde{y}_{0,T-1} = \text{new non-believers}_z &\begin{pmatrix} 0_{z,z} & 0_{z,T-z} \\ 0_{T-z,z} & J_{[0\bullet T-z-1,0\bullet T-z-1]} \end{pmatrix} \tilde{\psi}_{0,T-1} \qquad \text{[unlearn at } z\text{]} && \text{(C.69)} \\
+ \text{believers}_z &\begin{pmatrix} 0_{z,1} \\ J_{[0\bullet T-1-z,0]} \end{pmatrix} \tilde{\psi}_z && \text{(C.70)} \\
+ \text{new non-believers}_{z+1} &\begin{pmatrix} 0_{z+1,z+1} & 0_{z+1,T-z-1} \\ 0_{T-z-1,z+1} & J_{[0\bullet T-z-2,0\bullet T-z-2]} \end{pmatrix} \tilde{\psi}_{0,T-1} \quad \text{[unlearn at } z+1\text{]} && \text{(C.71)} \\
+ \text{believers}_{z+1} &\begin{pmatrix} 0_{z+1,1} \\ J_{[0\bullet T-2-z,0]} \end{pmatrix} \tilde{\psi}_{z+1} && \text{(C.72)} \\
&\vdots && \text{(C.73)} \\
+\text{new non-believers}_{T-1} &\begin{pmatrix} 0_{T-1,T-1} & 0_{T-1,1} \\ 0_{1,T-1} & J_{[0\bullet 0,0\bullet 0]} \end{pmatrix} \tilde{\psi}_{0,T-1} \Big\} && \text{(C.74)} \\
+ \text{believers}_{T-1} &\begin{pmatrix} 0_{T-1,1} \\ J_{[0\bullet 0,0]} \end{pmatrix} \tilde{\psi}_{T-1} \qquad \text{[unlearn at } T-1\text{]} && \text{(C.75)}
\end{aligned}
$$

We can define the sticky-information Jacobian W such that the outcome effect of the second wave is:

$$
\tilde{y}_{0,T-1} = W\tilde{\psi}_{0,T-1} \tag{C.76}
$$

W can be constructed recursively using:

$$
\begin{aligned}
W^1_{[k\bullet T-1,k\bullet T-1]} &= W^1_{[k\bullet T-1,k\bullet T-1]} + \text{new non-believers}_k J_{[0\bullet T-1-k,0\bullet T-1-k]}, \qquad k = s, \ldots, T-1 && \text{(C.77)} \\
W^2_{[k\bullet T-1,k]} &= \text{believers}_k J_{[0\bullet T-1-k]}, \qquad k = s, \ldots, T-1 && \text{(C.78)} \\
W &= W^1 + W^2 && \text{(C.79)}
\end{aligned}
$$

Combining both waves

The total effect of the two waves is

$$
\begin{aligned}
y_{0,T-1} &= S\psi^0_{0,T-1} + W\tilde{\psi}_{0,T-1} && \text{(C.80)} \\
&= (S-W)\psi^0_{0,T-1} && \text{(C.81)}
\end{aligned}
$$

The matrix $K = S - W$ governs the transmission of a fake news shock with path given by $\psi^0_{0,T-1}$.

Bibliography

Rodrigo Adao, Costas Arkolakis, and Federico Esposito. General equilibrium effects in space: Theory and measurement. Technical report, National Bureau of Economic Research, 2019a.

Rodrigo Adao, Michal Kolesár, and Eduardo Morales. Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010, 2019b.

Laura Alfaro, Manuel García-Santana, and Enrique Moral-Benito. On the direct and indirect real effects of credit supply shocks. *Journal of Financial Economics*, 139(3):895–921, 2021.

George Allayannis, Gregory W Brown, and Leora F Klapper. Capital structure and financial risk: Evidence from foreign debt use in east asia. *The Journal of Finance*, 58(6):2667–2710, 2003.

Mattias Almgren, José-Elías Gallegos, John Kramer, and Ricardo Lima. Monetary policy and liquidity constraints: Evidence from the euro area. *American Economic Journal: Macroeconomics*, 14(4): 309–340, 2022.

Martín Almuzara and Víctor Sancibrián. Micro responses to macro shocks. *FRB of New York Staff Report*, (1090), 2024.

Annette Alstadsæter, Niels Johannesen, and Gabriel Zucman. Who owns the wealth in tax havens? macro evidence and implications for global inequality. *Journal of Public Economics*, 162:89–100, 2018.

Mary Amiti and David E Weinstein. Exports and financial shocks. *The Quarterly Journal of Economics*, 126(4):1841–1877, 2011.

Mary Amiti and David E Weinstein. How much do idiosyncratic bank shocks affect investment? evidence from matched bank-firm loan data. *Journal of Political Economy*, 126(2):525–587, 2018.

Marıa Aristizábal-Ramırez and Christian Posso. Dynamics of corporate credit markets, employment and wages: Evidence from colombia. Technical report, Working paper, 2024.

Dmitry Arkhangelsky and Vasily Korovkin. On policy evaluation with aggregate time-series instruments. Technical report, CEPR Discussion Papers, 2024.

Borağan Aruoba and Thomas Drechsel. Identifying monetary policy shocks: A natural language approach. Technical report, National Bureau of Economic Research, 2024.

Adrien Auclert and Matthew Rognlie. Inequality and aggregate demand. Technical report, National Bureau of Economic Research, 2018.

Adrien Auclert, Matthew Rognlie, and Ludwig Straub. Micro jumps, macro humps: Monetary policy and business cycles in an estimated hank model. Technical report, National Bureau of Economic Research, 2020.

Adrien Auclert, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. *Econometrica*, 89(5):2375–2408, 2021.

Alan Auerbach and Yuriy Gorodnichenko. How powerful are fiscal multipliers in recessions? *NBER Reporter*, (2):21–24, 2015.

Alan Auerbach, Yuriy Gorodnichenko, and Daniel Murphy. Local fiscal multipliers and fiscal spillovers in the usa. *IMF Economic Review*, 68:195–229, 2020.

David Autor, David Dorn, and Gordon H Hanson. On the persistence of the china shock. Technical report, National Bureau of Economic Research, 2021.

David H Autor, David Dorn, and Gordon H Hanson. The china syndrome: Local labor market effects of import competition in the united states. *American economic review*, 103(6):2121–2168, 2013.

Rüdiger Bachmann and Eric R Sims. Uncertainty and the effectiveness of fiscal policy in the us and europe. *American Economic Review*, 102(3):60–65, 2012.

Jushan Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, 2009.

Laurence Ball. Comment on 'inflation and activity' by olivier blanchard, eugenio cerutti and lawrence summers. *ECB, Inflation and Unemployment in Europe, Frankfurt am Main, October*, pages 47–52, 2015.

Régis Barnichon and Geert Mesters. A sufficient statistics approach for macro policy. *American Economic Review*, 113(11):2809–2845, 2023.

Paul Beaudry, Dana Galizia, and Franck Portier. Reconciling hayek's and keynes' views of recessions. *The Review of Economic Studies*, 85(1):119–156, 2018.

Gianluca Benigno and Luca Fornaro. Stagnation traps. *The Review of Economic Studies*, 85(3):1425–1470, 2018.

Martin Beraja and Christian K Wolf. Demand composition and the strength of recoveries. 2022.

Olivier Blanchard, Eugenio Cerutti, and Lawrence Summers. Inflation and activity–two explorations and their monetary policy implications. Technical report, National Bureau of Economic Research, 2015.

Olivier Jean Blanchard and Lawrence F Katz. Regional evolutions. *Brookings Papers on Economic Activity*, 1992(1):1–75, 1992.

Kirill Borusyak and Peter Hull. Nonrandom exposure to exogenous shocks. *Econometrica*, 91(6): 2155–2185, 2023.

Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213, 2022.

Nina Boyarchenko and Leonardo Elias. The global credit cycle. *FRB of New York Staff Report*, (1094), 2024.

Martin Brown, Steven Ongena, and Pinar Yeşin. Foreign currency borrowing by small firms in the transition economies. *Journal of Financial Intermediation*, 20(3):285–302, 2011.

Martin Brown, Steven Ongena, and Pinar Yeşin. Information asymmetry and foreign currency borrowing by small firms. *Comparative Economic Studies*, 56(1):110–131, 2014.

Valentina Bruno and Hyun Song Shin. Dollar and exports. *The Review of Financial Studies*, 36(8): 2963–2996, 2023.

Andrea Caggese and Vicente Cuñat. Financing constraints, firm dynamics, export decisions, and aggregate productivity. *Review of Economic Dynamics*, 16(1):177–193, 2013.

Guillermo Calvo. On dollarization. *Economics of transition*, 10(2):393–403, 2002.

Fabio Canova. Should we trust cross-sectional multiplier estimates? *Journal of Applied Econometrics*, 39 (4):589–606, 2024.

Fabio Canova and Evi Pappa. Fiscal policy, pricing frictions and monetary accommodation. *Economic Policy*, 26(1):555–598, 2011.

Tomás E Caravello, Alisdair McKay, and Christian Wolf. Evaluating policy counterfactuals: A var-plus approach. Technical report, National Bureau of Economic Research, 2024.

Gerald Carlino and Robert DeFina. The differential regional effects of monetary policy. *Review of economics and statistics*, 80(4):572–587, 1998.

Christopher D Carroll, Edmund Crawley, Jiri Slacalek, Kiichi Tokuoka, and Matthew N White. Sticky expectations and consumption dynamics. *American economic journal: macroeconomics*, 12(3):40–76, 2020.

Gabriel Chodorow-Reich. The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis. *The Quarterly Journal of Economics*, 129(1):1–59, 2014.

Gabriel Chodorow-Reich. Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy*, 11(2):1–34, 2019.

Gabriel Chodorow-Reich. Regional data in macroeconomics: Some advice for practitioners. *Journal of Economic Dynamics and Control*, 115:103875, 2020.

Gabriel Chodorow-Reich, Laura Feiveson, Zachary Liscow, and William Gui Woolston. Does state fiscal relief during recessions increase employment? evidence from the american recovery and reinvestment act. *American Economic Journal: Economic Policy*, 4(3):118–145, 2012.

Gabriel Chodorow-Reich, Plamen T Nenov, and Alp Simsek. Stock market wealth and the real economy: A local labor market approach. *American Economic Review*, 111(5):1613–1657, 2021.

Davin Chor and Kalina Manova. Off the cliff and back? credit conditions and international trade during the global financial crisis. *Journal of international economics*, 87(1):117–133, 2012.

Lawrence Christiano, Martin Eichenbaum, and Sergio Rebelo. When is the government spending multiplier large? *Journal of Political Economy*, 119(1):78–121, 2011.

James S Cloyne, Òscar Jordà, and Alan M Taylor. Decomposing the fiscal multiplier. Technical report, National Bureau of Economic Research, 2020.

Olivier Coibion. Are the effects of monetary policy shocks big or small? *American Economic Journal: Macroeconomics*, 4(2):1–32, 2012.

Hans Degryse, Olena Havrylchyk, Emilia Jurzyk, and Sylwester Kozak. Foreign bank entry, credit allocation and lending rates in emerging markets: Empirical evidence from poland. *Journal of Banking & Finance*, 36(11):2949–2959, 2012.

Marco Del Negro, Marc P Giannoni, and Christina Patterson. The forward guidance puzzle. *Journal of Political Economy Macroeconomics*, 1(1):43–79, 2023.

J Bradford DeLong and Lawrence H Summers. Fiscal policy in a depressed economy. *Brookings Papers on Economic Activity*, 43(1):233–297, 2012.

Yuliya Demyanyk, Elena Loutskina, and Daniel Murphy. Fiscal stimulus and consumer debt. *Review of Economics and Statistics*, 101(4):728–741, 2019.

Bill Dupor and Rodrigo Guerrero. Local and aggregate fiscal policy multipliers. *Journal of Monetary Economics*, 92:16–30, 2017.

Bill Dupor, Marios Karabarbounis, Marianna Kudlyak, and M Saif Mehkari. Regional consumption responses and the aggregate fiscal multiplier. *Review of Economic Studies*, 90(6):2982–3021, 2023.

Stéphane Dupraz, Emi Nakamura, and Jón Steinsson. A plucking model of business cycles. 2023.

Leonardo Elias. Capital flows and the real effects of corporate rollover risk. Technical report, Working paper, Federal Reserve Bank of New York, 2021.

George W Evans and Garey Ramey. Expectation calculation and macroeconomic dynamics. *The American Economic Review*, pages 207–224, 1992.

Emmanuel Farhi and Iván Werning. Fiscal multipliers: Liquidity traps and currency unions. In *Handbook of macroeconomics*, volume 2, pages 2417–2492. Elsevier, 2016.

Stefano Federico, Fadi Hassan, and Veronica Rappoport. Trade shocks and credit reallocation. Technical report, National Bureau of Economic Research, 2023.

John Fernald, Robert Inklaar, and Dimitrije Ruzic. The productivity slowdown in advanced economies: common shocks or common trends? *Review of Income and Wealth*, 71(1):e12690, 2025.

Andrew J Fieldhouse, David Munro, Christoffer Koch, and Sean Howard. The emergence of a uniform business cycle in the united states: Evidence from new claims-based unemployment data. *Brookings Papers on Economic Activity*, 2004(1):265–319, 2024.

Andrew T Foerster, Pierre-Daniel G Sarte, and Mark W Watson. Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy*, 119(1):1–38, 2011.

Andrew T Foerster, Andreas Hornstein, Pierre-Daniel G Sarte, and Mark W Watson. Aggregate implications of changing sectoral trends. *Journal of Political Economy*, 130(12):3286–3333, 2022.

Milton Friedman. The “plucking model” of business fluctuations revisited. *Economic Inquiry*, 31(2): 171–177, 1993.

Jordi Galí, Frank Smets, and Rafael Wouters. Slow recoveries: A structural interpretation. *Journal of Money, Credit and Banking*, 44:9–30, 2012.

Arturo Galindo, Ugo Panizza, and Fabio Schiantarelli. Debt composition and balance sheet effects of currency depreciation: a summary of the micro evidence. *Emerging Markets Review*, 4(4):330–339, 2003.

Vaishali Garga and Sanjay R Singh. Output hysteresis and optimal monetary policy. *Journal of Monetary Economics*, 117:871–886, 2021.

Patricia Gil, Justin Holz, John A. List, Andrew Simon, and Alejandro Zentner. Toward an understanding of tax amnesty take-up: Evidence from a natural field experiment. *Journal of Public Economics*, 239: 105245, 2024. doi: 10.1016/j.jpubeco.2024.105245.

Simon Gilchrist and Charles P Himmelberg. Evidence on the role of cash flow for investment. *Journal of monetary Economics*, 36(3):541–572, 1995.

Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624, 2020.

Martın Gonzalez-Rozada, Pablo Andrés Neumeyer, Alejandra Clemente, et al. The elasticity of substitution in demand for non-tradable goods in latin america. case study: Argentina. *Universidad T. Di Tella Working Paper*, 27, 2003.

Adam Guren, Alisdair McKay, Emi Nakamura, and Jón Steinsson. What do we learn from cross-regional empirical estimates in macroeconomics? *NBER Macroeconomics Annual*, 35(1):175–223, 2021a.

Adam M Guren, Alisdair McKay, Emi Nakamura, and Jón Steinsson. Housing wealth effects: The long view. *The Review of Economic Studies*, 88(2):669–707, 2021b.

Lukas Hack, Klodiana Istrefi, and Matthias Meier. Identification of systematic monetary policy. 2023.

Bryan Hardy. Foreign currency borrowing, balance sheet shocks and real outcomes. *Balance Sheet Shocks and Real Outcomes (November 22, 2018). BIS Working Paper*, 1(758), 2018.

Bryan Hardy. Foreign currency borrowing, balance sheet shocks, and real outcomes. *Journal of International Money and Finance*, 139:102969, 2023.

Juan Herreño. Aggregating the effect of bank credit supply shocks on firms. *Working Paper*, 2023.

Juan Herreño and Mathieu Pedemonte. The geographic effects of monetary policy shocks. *Review of Economics and Statistics*, pages 1–46, 2025.

Christopher L House and Matthew D Shapiro. Temporary investment tax incentives: Theory with evidence from bonus depreciation. *American Economic Review*, 98(3):737–768, 2008.

Greg Howard, Robert Martin, and Beth Anne Wilson. Are recoveries from banking and financial crises really so different? 2011.

Guido W Imbens and Michal Kolesar. Robust standard errors in small samples: Some practical advice. *Review of Economics and Statistics*, 98(4):701–712, 2016.

Gabriel Jiménez, Atif Mian, José-Luis Peydró, and Jesús Saurina. The real effects of the bank lending channel. *Journal of Monetary Economics*, 115:162–179, 2020.

Òscar Jordà. Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182, 2005.

Òscar Jordà, Moritz Schularick, and Alan M Taylor. When credit bites back. *Journal of money, credit and banking*, 45(s2):3–28, 2013.

Sebnem Kalemli-Ozcan, Herman Kamil, and Carolina Villegas-Sanchez. What hinders investment in the aftermath of financial crises: Insolvent firms or illiquid banks? *Review of Economics and Statistics*, 98 (4):756–769, 2016.

Herman Kamil. *How do exchange rate regimes affect firms' incentives to hedge currency risk? Micro evidence for Latin America*. International Monetary Fund, 2012.

Anil K Kashyap and Jeremy C Stein. What do a million observations on banks say about the transmission of monetary policy? *American Economic Review*, 90(3):407–428, 2000.

Akihisa Kato, Wataru Miyamoto, Thuy Lan Nguyen, and Dmitriy Sergeyev. The effects of tax changes at the zero lower bound: Evidence from japan. *AEA Papers and Proceedings*, 108:513–518, 2018.

Patrick J Kehoe, Virgiliu Midrigan, and Elena Pastorino. Evolution of modern business cycle models: Accounting for the great recession. *Journal of Economic Perspectives*, 32(3):141–166, 2018.

Asim Ijaz Khwaja and Atif Mian. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review*, 98(4):1413–42, 2008.

David Kohn, Fernando Leibovici, and Michal Szkup. Financial frictions and new exporter dynamics. *International economic review*, 57(2):453–486, 2016.

David Kohn, Fernando Leibovici, and Michal Szkup. No credit, no gain: Trade liberalization dynamics, production inputs, and financial development. *International Economic Review*, 64(2):809–836, 2023.

Julian Kozlowski, Laura Veldkamp, and Venky Venkateswaran. The tail that keeps the riskless rate low. *NBER Macroeconomics Annual*, 33(1):253–283, 2019.

Julian Kozlowski, Laura Veldkamp, and Venky Venkateswaran. The tail that wags the economy: Beliefs and persistent stagnation. *Journal of Political Economy*, 128(8):2839–2879, 2020.

Dominika Langenmayr. Voluntary disclosure of evaded taxes—increasing revenue, or increasing incentives to evade? *Journal of Public Economics*, 151:110–125, 2017.

Mr Eric Le Borgne and Ms Katherine Baer. *Tax amnesties: Theory, trends, and some alternatives*. International Monetary Fund, 2008.

Edward E Leamer. Why are some recoveries short and others long? Technical report, National Bureau of Economic Research, 2021.

Eric M Leeper, Nora Traum, and Todd B Walker. Clearing up the fiscal multiplier morass. *American Economic Review*, 107(8):2409–2454, 2017.

Fernando Leibovici. Financial development and international trade. *Journal of Political Economy*, 129 (12):3405–3446, 2021.

Arjan Lejour, Wouter Leenders, Simon Rabate, and Maarten van't Riet. Offshore tax evasion and wealth inequality: Evidence from a tax amnesty in the netherlands. *Journal of Public Economics*, 2022.

Andrei A Levchenko, Logan T Lewis, and Linda L Tesar. The collapse of international trade during the 2008–09 crisis: in search of the smoking gun. *IMF Economic review*, 58(2):214–253, 2010.

Juliana Londoño-Vélez and Dario Tortarolo. Revealing 21% of gdp in hidden assets: Evidence from argentina's tax amnesties. Technical report, UNU WIDER Working Paper 103, 2022.

Juliana Londoño-Vélez, Dario Tortarolo, et al. Revealing 21 per cent of gdp in hidden assets: Evidence from argentina's tax amnesties. Technical report, World Institute for Development Economic Research (UNU-WIDER), 2022.

Kalina Manova. Credit constraints, equity market liberalizations and international trade. *Journal of International Economics*, 76(1):33–47, 2008. ISSN 0022-1996. doi: https://doi.org/10.1016/j.jinteco.2008.03.008. URL https://www.sciencedirect.com/science/article/pii/S0022199608000305.

Kalina Manova. Credit Constraints, Heterogeneous Firms, and International Trade. *The Review of Economic Studies*, 80(2):711–744, 11 2012. ISSN 0034-6527. doi: 10.1093/restud/rds036. URL https://doi.org/10.1093/restud/rds036.

Ed Manuel. Identification in general equilibrium: The missing intercept problem revisited. Technical report, Working Paper, 2025.

Christian Matthes, Naoya Nagasaka, and Felipe Schwartzman. Estimating the missing intercept. 2024.

Alisdair McKay and Johannes F Wieland. Forward guidance and durable goods demand. *American Economic Review: Insights*, 4(1):106–122, 2022.

Alisdair McKay and Wolf. What can time-series regressions tell us about policy counterfactuals? *Econometrica*, 91(5):1695–1725, 2023.

Alisdair McKay, Emi Nakamura, and Jón Steinsson. The power of forward guidance revisited. *American Economic Review*, 106(10):3133–3158, 2016.

Alisdair McKay, Emi Nakamura, and Jón Steinsson. The discounted euler equation: A note. *Economica*, 84(336):820–831, 2017.

Atif Mian and Amir Sufi. What explains the 2007–2009 drop in employment? *Econometrica*, 82(6): 2197–2223, 2014.

Silvia Miranda-Agrippino and Hélène Rey. The global financial cycle. In *Handbook of international economics*, volume 6, pages 1–43. Elsevier, 2022.

Nada Mora. The bank lending channel in a partially dollarized economy. *Journal of Applied Economics*, 16(1):121–151, 2013.

Mirabelle Muûls. Exporters, importers and credit constraints. *Journal of International Economics*, 95(2): 333–343, 2015.

Emi Nakamura and Jón Steinsson. Fiscal stimulus in a monetary union: Evidence from us regions. *American Economic Review*, 104(3):753–792, 2014.

Emi Nakamura and Jón Steinsson. Identification in macroeconomics. *Journal of Economic Perspectives*, 32(3):59–86, 2018.

Maurice Obstfeld and Haonan Zhou. The global dollar cycle. *Brookings Papers on Economic Activity*, 2022(2):361–447, 2022.

Jacob Orchard, Valerie A Ramey, and Johannes F Wieland. Micro mpcs and macro counterfactuals: the case of the 2008 rebates. Technical report, National Bureau of Economic Research, 2023.

Pablo Ottonello and Thomas Winberry. Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, 88(6):2473–2502, 2020.

Daniel Paravisini. Local bank financial constraints and firm access to external finance. *The Journal of Finance*, 63(5):2161–2193, 2008.

Daniel Paravisini, Veronica Rappoport, and Philipp Schnabl. Specialization in bank lending: Evidence from exporting firms. Technical report, National Bureau of Economic Research, 2015a.

Daniel Paravisini, Veronica Rappoport, Philipp Schnabl, and Daniel Wolfenzon. Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. *The review of economic studies*, 82 (1):333–359, 2015b.

Daniel Paravisini, Veronica Rappoport, Philipp Schnabl, and Daniel Wolfenzon. Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. *The Review of Economic Studies*, 82 (1):333–359, 2015c.

Jonathan A Parker, Nicholas S Souleles, David S Johnson, and Robert McClelland. Consumer spending and the economic stimulus payments of 2008. *American economic review*, 103(6):2530–2553, 2013.

Steven Pennings. Cross-region transfer multipliers in a monetary union: Evidence from social security and stimulus payments. *American Economic Review*, 111(5):1689–1719, 2021.

Steven Pennings. Locally financed and outside financed regional fiscal multipliers. *Economics Letters*, 213:110389, 2022.

José-Luis Peydró, Andrea Polo, and Enrico Sette. Monetary policy at work: Security and credit application registers evidence. *Journal of Financial Economics*, 140(3):789–814, 2021.

Valerie A Ramey. Macroeconomic shocks and their propagation. *Handbook of macroeconomics*, 2:71–162, 2016.

Valerie A Ramey. Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives*, 33(2):89–114, 2019.

Valerie A. Ramey. Discussion of Guren, McKay, Nakamura, Steinsson "What Do We Learn from Cross-Sectional Empirical Estimates in Macroeconomics?". *NBER Macroeconomics Annual*, 35, 2020.

Valerie A Ramey and Sarah Zubairy. Government spending multipliers in good times and in bad: Evidence from us historical data. *Journal of Political Economy*, 126(2):850–901, 2018.

Dave Reifschneider, William Wascher, and David Wilcox. Aggregate supply in the united states: recent developments and implications for the conduct of monetary policy. *IMF Economic Review*, 63(1): 71–109, 2015.

Helene Rey. The international credit channel and monetary autonomy. *Mundell-Fleming Lecture, International Monetary Fund*, 13, 2014.

Daniel Riera-Crichton, Carlos A Vegh, and Guillermo Vuletin. Fiscal solvency and procyclical fiscal policy. *Journal of Development Economics*, 117:305–319, 2015.

Matthew Rognlie, Andrei Shleifer, and Alp Simsek. Investment hangover and the great recession. *American Economic Journal: Macroeconomics*, 10(2):113–153, 2018.

Christina D Romer and David H Romer. A new measure of monetary shocks: Derivation and implications. *American economic review*, 94(4):1055–1084, 2004.

Christina D Romer and David H Romer. Presidential address: Does monetary policy matter? the narrative approach after 35 years. *American Economic Review*, 113(6):1395–1423, 2023.

Andrés Pablo Sarto. Recovering macro elasticities from regional data. 2024.

Philipp Schnabl. The international transmission of bank liquidity shocks: Evidence from an emerging market. *The Journal of Finance*, 67(3):897–932, 2012.

Christopher A Sims and Tao Zha. Were there regime switches in us monetary policy? *American Economic Review*, 96(1):54–81, 2006.

James H Stock and Mark W Watson. Business cycle fluctuations in us macroeconomic time series. *Handbook of macroeconomics*, 1:3–64, 1999.

Alonso Villacorta, Lucciano Villacorta, and Bryan Gutierrez. (unobserved) heterogeneity in the bank lending channel: Accounting for bank-firm interactions and specialization. *Available at SSRN 4608656*, 2023.

Johannes F Wieland and Mu-Jeung Yang. Financial dampening. *Journal of Money, Credit and Banking*, 52(1):79–113, 2020.

Christian Wolf. The missing intercept in cross-regional regressions. Technical report, Working Paper., 2019.

Christian Wolf. Fiscal stimulus and the systematic response of monetary policy. In *AEA Papers and Proceedings*, volume 113, pages 388–393, 2023a.

Christian Wolf. The missing intercept: A demand equivalence approach. *American Economic Review*, 113(8):2232–2269, 2023b.

Danny Yagan. Employment hysteresis from the great recession. *Journal of Political Economy*, 127(5): 2505–2558, 2019.

Owen Zidar. Tax cuts for whom? heterogeneous effects of income tax changes on growth and employment. *Journal of Political Economy*, 127(3):1437–1472, 2019.

Eric Zwick and James Mahon. Tax policy and heterogeneous investment behavior. *American Economic Review*, 107(1):217–248, 2017.