Title

Essays in Household Finance and Empirical Macroeconomics

ISBN

9798288862939

Author

Boctor, Valerie

Publication Date

2025-05-15

Peer reviewed|Thesis/dissertation

Essays on Household Finance and Empirical Macroeconomics

By

Valerie Boctor

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jón Steinsson, Co-Chair
Professor Benjamin Schoefer, Co-Chair
Professor David Romer
Professor Amir Kermani

Spring 2025

Abstract

Essays on Household Finance and Empirical Macroeconomics

by

Valerie Boctor

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Jón Steinsson, Co-Chair

Professor Benjamin Schoefer, Co-Chair

This dissertation is composed of three chapters that cover a broad range of policy-relevant topics in household finance and empirical macroeconomics. The first chapter, titled "Mortgage Forbearance and Financial Distress in the Long Run," the causal effects of mortgage forbearance under the Coronavirus Aid, Relief, and Economic Security (CARES) Act on household financial stability. Leveraging quasi-random variation in mortgage servicers' forbearance provision, I identify significant reductions in mortgage delinquency rates—up to 5 percentage points—and foreclosure rates by 1 percentage point, persisting three years post-forbearance. Additionally, the program had beneficial spillover effects on revolving credit stability, reducing credit card delinquencies by 2 percentage points and utilization rates by roughly 15 percentage points relative to the pre-pandemic period. Upon exiting forbearance, borrowers not only avoided financial 'rebound effects,' but also sustained improved financial stability for more than two years following the policy's implementation.

The second chapter, coauthored with Yuriy Gorodnichenko, Olivier Coibion, and Michael Weber, explores how survey design affects the measurement of household macroeconomic expectations. We show that responses—particularly inflation expectations—are highly sensitive to question wording, even within a single 15-minute survey. These discrepancies raise concerns for policymakers relying on such data and underscore the need for further research to understand the underlying sources of measurement error.

The third chapter, written in collaboration with Ryan Banerjee, Fabrizio Zampolli, and Aaron Mehrotra at the Bank for International Settlements, examines how the inflationary effects of fiscal deficits depend on a country's prevailing fiscal-monetary policy regime. Using a panel of advanced economies over four decades, we show that inflation-at-risk rises significantly under fiscally-led regimes. We calibrate an inflation-at-risk model and find that post-COVID-19 fiscal stimulus outcomes are broadly consistent with the predictions of fiscally-led regimes, highlighting the importance of institutional context in shaping macroeconomic outcomes.

Acknowledgements

Beyond economics, completing this dissertation taught me profound lessons about resilience, staying true, growing a spine, and protecting what I love. First and foremost, I want to thank my mom, Eman Tadros, and dad, Ragi Boctor, for anchoring me and believing in me at every step in the journey. I am deeply indebted to the many role models in my family who embody the kind of steely courage and risk-taking spirit that inspired me—especially in the later years of my PhD. I would also like to thank my siblings, Lorine and Jizelle, for being my natural best friends and reminding me that, at the end of the day, I am just a little sister.

I am also sincerely grateful to my co-chairs, Jón Steinsson and Benjamin Schoefer, for their invaluable support and feedback on my dissertation—and for bearing witness to many spells of chaos that went into its making (and likely tortured their eyeballs in the process). I would also like to thank David Romer for being a steady source of wisdom throughout the writing process, Amir Kermani for sharing his expertise on mortgage markets, and Yuriy Gorodnichenko for providing initial research support.

To the mentors who breathed life into my young career and believed in me long before I believed in myself, I'm truly grateful. Raza Habib Raja from Syracuse University and Paul Reymann and Andrea Shearin from the Office of the Comptroller of the Currency had an especially profound impact.

A very special thanks go to the faculty at the Syracuse University Economics Department, Gary V. Engelhardt, Perry Singleton, Jan Ondrich, and Michael J. Wasylenko, who taught me never to assume I'm a slam dunk, but to shoot my shot anyway.

Thanks also to my wonderful coauthors at the Bank for International Settlements, Ryan Banerjee, Aaron Mehrotra, and Fabrizio Zampolli, for their collaboration and insight. I am also grateful to my key sources of research funding: the California Policy Lab, the Berkeley Opportunity Lab, and the BB90 Fund whose generous support made this work possible.

To my friends and cohort members: thank you for the many laughs, hikes, LaTex templates, and Python support. I will always admire how brilliant, kind, and diligent you are and wish you all the best with your future endeavors.

Finally, to my grandmother-whose will to survive compels me to believe in miracles—this work is for you.

Introduction

Understanding how macroeconomic policy shapes household behavior has become increasingly vital in the wake of recent global shocks. This dissertation investigates the interplay between household finance, policy design, and macroeconomic risk through three distinct but thematically connected chapters. Each chapter addresses a timely and policy-relevant question using empirical methods rooted in applied micro techniques and core macroeconomic questions.

The first chapter focuses on mortgage forbearance policy during the COVID-19 pandemic—a large-scale natural experiment in household financial relief. Leveraging administrative credit bureau data covering 500,000 consumers and exploiting quasi-random differences in mortgage servicers' implementation of the CARES Act, I identify the causal effects of forbearance on long-term household financial outcomes. The analysis reveals that mortgage forbearance not only substantially reduced delinquency and foreclosure rates in the short term but also promoted sustained financial stability well beyond the policy's end. Additionally, I document positive spillover effects on revolving credit, offering a broader view of how temporary debt relief can influence household balance sheets and liquidity management.

The second chapter, coauthored with Yuriy Gorodnichenko, Olivier Coibion, and Michael Weber, examines the reliability of survey-based measures of household macroeconomic expectations. We show that small changes in survey wording can significantly alter reported inflation expectations, even when the questions are presented in the same questionnaire in a short window. These findings highlight the fragility of widely used survey metrics and call for careful design and interpretation when such data are used to guide or evaluate monetary policy.

The third chapter, developed in collaboration with researchers at the Bank for International Settlements, explores how the inflationary consequences of fiscal policy depend on the prevailing regime of fiscal and monetary coordination. Using a panel of advanced economies over four decades, we find that fiscal expansions pose greater inflationary risk under fiscally-led regimes, where monetary authorities are constrained in offsetting fiscal pressures. Through an inflation-at-risk framework, we interpret the post-pandemic inflation experience through the lens of regime classification, showing that the inflation dynamics of recent years align with patterns observed in fiscally dominant environments.

Together, these chapters contribute to a deeper understanding of how policy interacts with household behavior and macroeconomic risk, offering insights with direct implications for the design of financial relief programs, the interpretation of expectations data, and the coordination of fiscal and monetary policy.

Chapter 1: Mortgage Forbearance and Financial Stability in the Long Run*

Valerie Boctor†

May 17, 2025

Abstract

Mortgage relief programs are crucial for distressed households during economic downturns, but their long-term effects remain underexplored. This paper offers new micro-level evidence on the long-term efficacy of mortgage payment pauses, or forbearance, in mitigating financial distress during and after the COVID-19 pandemic. Using data from 500,000 consumer credit reports, I study the causal effects of mortgage forbearance under the Coronavirus Aid, Relief, and Economic Security (CARES) Act on household financial stability. Leveraging quasi-random variation in mortgage servicers' forbearance provision, I identify significant reductions in mortgage delinquency rates—up to 5 percentage points—and foreclosure rates by 1 percentage point, persisting three years post-forbearance. Additionally, the program had beneficial spillover effects on revolving credit stability, reducing credit card delinquencies by 2 percentage points and utilization rates by roughly 15 percentage points relative to the pre-pandemic period. Upon exiting forbearance, borrowers not only avoided financial 'rebound effects,' but also sustained improved financial stability for more than two years following the policy's implementation.

JEL Codes: D12; R28; G21
Keywords: Mortgage forbearance, targeted debt relief, housing policy, household financial stability

*This research is supported by the Berkeley Opportunity Lab Place-Based Policy Initiative, the California Policy Lab, and the BB90 Fund for Monetary Economics.

†University of California, Berkeley

1 Introduction

Financial stability is a key measure of households' economic well-being and predictor of social mobility. During the COVID-19 pandemic, roughly 40% of households reported facing serious financial stress, with debt obligations being the most commonly reported reason (Calfas, 2021). This paper examines the long-term impacts of federal mortgage relief programs during the Coronavirus pandemic. Through a provision of the Coronavirus Aid, Relief and Economic Stimulus (CARES) Act, mortgage servicers were mandated pause the mortgage payments of eligible homeowners for up to 12 months, as long as they self-attested to a COVID-related financial hardship.[1] Using the credit reports of 500,000 eligible homeowners, I leverage quasi-random variation in mortgage servicers' forbearance provision to estimate the program's causal impact on mortgage and revolving credit stability up to three years after CARES was implemented. The results reveal significant reductions in mortgage delinquency rates—up to 5 percentage points—and foreclosure rates by 1 percentage point, persisting three years post-forbearance. Additionally, the program had beneficial spillover effects on revolving credit stability, reducing credit card delinquencies by 2 percentage points and utilization rates by roughly 15 percentage points relative to the pre-pandemic period. Upon exiting forbearance, borrowers not only avoided financial 'rebound effects,' but also sustained improved financial stability for more than two years following the policy's implementation.

Because mortgage payments make up roughly one third of American households' monthly income, the risk of falling behind on these payments during recessions is both high and costly, with potential aggregate consequences.[2] The mortgage forbearance provision of the CARES Act was part of a concerted policy effort to prevent a wave of financial distress due to the massive employment disruptions as a result of lockdowns. At the height of the program's popularity in 2020Q2, approximately 8% of mortgaged homeowners were in forbearance, covering $1.4 trillion of loan value, and allowing borrowers to miss about $31 billion in payments.(Cherry et al., 2021). Carrying forward important lessons from the Great Recession, the housing policy response to COVID, including the forbearance provision of the CARES Act, featured generous terms, broad reach and an immediate impact for a large swath of American households. The forbearance provision of the CARES Act, in particular, was designed to improve upon the Great Recession era Home Affordable Refinance Program (HARP) and Home Affordable Modification Program (HAMP) by significantly reducing paperwork requirements and covering a much larger portion of households, placing the burden of the mandate on servicers to provide relief. From an optimal policy perspective, the unprecedentedly generous terms of the CARES forbearance program represent an important trade-off between protecting the balance sheets of servicers and lenders against the financial stability of households both at short and long horizons. This paper informs policymakers on the overall efficacy of generous countercyclical housing relief by showing that CARES forbearance was highly effective for participating households and did not lead to aggregate disruptions even as households exited forbearance in waves.

[1] In some cases, forbearance could be extended for a total of 18 months.

[2] A single missed mortgage payment costs the average borrower approximately $130 and leads to significant credit score reductions. Each additional delinquency leads to increased marginal costs and credit score reductions. In most states, missing 3 payments in a row triggers the foreclosure process

2 Institutional Background on Mortgage Forbearance

2.1 Mechanics of Loss Mitigation in Mortgage Markets

Mortgage forbearance is a short-term pause or reduction in payments designed to assist borrowers with temporarily financial hardship. Mortgage servicers, which act as clearinghouses between borrowers and investors, are primarily responsible for maintaining and recording loan performance, determining loss mitigation plans– including forbearance– when needed, and remitting payments to investors each month. The mortgage forbearance process typically begins when either the borrower contacts their servicer to request relief, or the servicer reaches out to borrower after flagging an underperforming loan. In principle, any borrower can apply for forbearance, regardless of whether their mortgage is backed by a government agency or private entity. To avoid straining their liquidity, mortgage servicers aim to limit forbearance periods, usually to three months, and screen borrowers carefully before recommending forbearance over other options, such as refinancing or selling the home. Forbearance exit strategies, and the implied levels of relief, vary, ranging from immediate payments reinstatements (i.e., balloon payments) at the more conservative end, to more generous, gradual repayment plans. The most common exit strategy from forbearance is reinstatement. Other exit plans include deferrals to maturity, refinancing, permanent loan modification, selling the home, or in severe distress cases, foreclosure. Ideal candidates for forbearance might include a tech worker furloughed by her firm, or a construction worker experiencing a one-off medical emergency.

2.2 The CARES Act

To alleviate the pandemic's economic side effects, policymakers unleashed an array of expansive policy measures targeted to businesses and households impacted by the lockdown. The CARES Act, introduced on March 27, 2020, represented the largest economic stimulus package in U.S. history, injecting \$2.2 billion of relief. Key components of the law included stimulus checks of \$1,200 per adult and \$500 per child for eligible taxpayers[3], expanded unemployment insurance (UI), small business loans, eviction and foreclosure moratoria, and forbearances on federal student loans and mortgages.[4]

The mortgage forbearance provision (Section 4022) of the CARES Act aimed to cover gaps that remained in the financial safety net after accounting for CARES and subsequent emergency relief policies. In general and in particular during economic downturns such as COVID, UI serves as the primary financial relief option for workers struggling to meet housing payment due to lost income. In response to the COVID-19 crisis, UI was enhanced with weekly supplements, expanded eligibility to self-employed, gig, and independent workers, and extended maximum UI duration by 53 weeks beyond each state's regular benefits.[5] Even with these expansions, some

[3]Eligibility for stimulus checks required income of less than \$75,000 for single filers or less than \$150,000 for joint filers.

[4]Unlike the mortgage forbearance provision, where borrowers needed to contact their servicers to opt in, the student loan relief was applied automatically to federal student loans.

[5]"Federal Pandemic Unemployment Compensation (FPUC) established weekly supplements on top of any state UI benefits for which recipients were eligible. Weekly supplements were available intermittently, and were set at \$600 between March and July 2020, \$300 in Lost Wages Assistance (LWA) in September and October 2020, and \$300 between January 2021 and September 6, 2021. Pandemic Unemployment Assistance (PUA) expanded eligibility of UI to self-employed workers, gig workers, independent workers, and others not previously eligible for UI or who were unable to work for a variety of COVID-related reasons. For example, workers could receive UI benefits if they were

unemployed workers may have found themselves unable to meet housing payment obligations due to caps on benefits, the lack of requirements for the funds to be spent on housing, and delays in when UI is actually received, particularly for households with little savings.(Calabria, 2023) Mortgage forbearance under the CARES Act was specifically designed to address these issues by providing (ideally prompt) liquidity to households by directly pausing mortgage payments.

In addition to mandating access to forbearance for federally backed homeowners, the provision waived interest, fees, or penalties on forborne payments and prohibited negative credit reporting from forbearance. The CARES forbearance mandate was silent on forbearance exit plans, although. Due to the policy's ambiguity towards exit plans, the Federal Housing Finance Agency as well as the CFPB and other agencies issued additional guidance on the breadth of options beyond lump sum balloon payments, such as deferrals to maturity modifications, and refinancing.(Mae, Mae; Consumer Financial Protection Bureau, nd) The CARES Act's generous forbearance terms meant that servicers were required to extend access to borrowers who might otherwise get screened out due to their hardship being considered either too minor or too severe. On the one hand, borrowers without substantial financial hardship were incentivized to opt in as a way to gain interest-free liquidity. On the other hand, severely distressed borrowers may have also opted in, in some cases falling further into distress or facing barriers to accessing more substantial servicer support after the forbearance period ended. On average, the program's wide reach did not lead to unintended financial consequences for participants, as evidenced by the policy's long term efficacy in terms of preventing mortgage distress and negative credit events.

2.2.1 Eligibility for Forbearance under CARES

In 2020, approximately 75% of mortgages in the US were federally backed and thus were covered by the CARES forbearance mandate. In general, a mortgage is determined to be federally backed at or shortly after origination, and it retains that status for the duration of the loan. The two broad classes of federally backed mortgages are those purchased by a Government Sponsored Enterprise (GSE), i.e., Fannie Mae or Freddie Mac, or insured by a government agency such as the Federal Housing Authority (FHA) or Veterans Affairs (VA). Because GSE-backed mortgages conform to strict underwriting guidelines and loan limits, borrowers typically have lower credit risk and moderate incomes. Conversely, FHA and VA mortgages are designed for low-income, high risk borrowers. (Cherry et al., 2021). These borrowers are allowed to make down payments as low at 3.5%. Under the CARES forbearance mandate, FHA and VA borrowers were twice as as likely as their GSE-backed counterparts to opt into the program (An et al., 2022).

2.2.2 Frictions to Forbearance Access

In principle, the CARES Act required servicers to provide forbearance to eligible borrowers in need, covering roughly 75% of mortgaged homeowners in the U.S. Despite the broad mandate, survey data from Fannie Mae during the pandemic indicated significant barriers to accessing forbearance due to informational and logistical frictions. In terms of providing relief to homeowners, the Fannie Mae Mortgage Lender Sentiment Survey found that nearly half of surveyed mortgage servicers (45%) cited difficulty in keeping up with investor policies

unable to work because of dependent care responsibilities, a COVID-19 illness in the family, or the health risk at work. Pandemic Emergency Unemployment Compensation (PEUC) extended the duration of federal UI benefits by 53 weeks for those who had exhausted their regular state benefits." (Ganong et al., 2022) around forbearance as their top or second-to-top challenge, followed by maintaining staffing capacity (reported by 34% of servicers) and training client-facing employees (29%).(Patane, 2021) The same survey highlighted challenges on the customer service front, with 34% of servicers listing the explanation of exit plans as a top or second-to-top challenge, followed by clarifying the implications of forbearance use (31%) and assessing readiness for forbearance exit (27%). Against this backdrop, the Consumer Financial Protection Bureau documented numerous instances in which servicers failed to comply with the forbearance mandate. Specific examples included servicers incorrectly sending borrowers collection or default notices, assessing fees, or initiating foreclosures for borrowers in forbearance; changing borrowers' scheduled payments without consent, failing to apply forbearance for requesting, eligible homeowners or applying forbearance to borrowers who did not request it, delayed processing times, and failure to establish forbearance exit plans.(Consumer Financial Protection Bureau, 2021)

On the homeowner side, these issues were reflected in a lack of awareness about forbearance and other mortgage relief policies, confusion around exit strategies and the financial implications of forbearance, and challenges in reaching servicers to initiate the process. Notably, data from the April - June 2020 waves of the Fannie Mae Housing Market Survey indicated that over 60% of homeowners were not aware of their COVID-19 mortgage relief options, including forbearance (Duncan, 2020). This widespread lack of awareness may explain why the forbearance rate during COVID peaked under 9% of households in 2020Q2.[6]According to reports from the Consumer Financial Protection Bureau (CFPB), information frictions also reflected in homeowners receiving incomplete or false information from servicers about their eligibility, fees associated with forbearance, and exit plans (such as being told balloon payments were required). (Kim et al., 2022; Consumer Financial Protection Bureau, 2021)Relatedly, Kim et al. (2022) and Cherry et al. (2021) document significant variation in the degree of servicer compliance and generosity in delivering forbearance. In Kim et al. (2022), up to one-third of VA and FHA borrowers become delinquent while never receiving forbearance during the pandemic, representing a missed opportunity for mortgage payment relief. They document key frictions including the servicers' idiosyncratic levels of generosity, capitalization and staffing issues, liquidity constraints, regulatory pressures acting on servicers of differing size, organizational form, and varying technological capacities. Taken together, these factors facilitated or hampered servicers willingness to provide borrowers with access to forbearance under the mandate. Cherry et al. (2021) corroborate the presence of liquidity- and organizational form-related frictions by noting that non-bank financial intermediaries (NBFI) and liquidity-constrained servicers had a lower likelihood of providing forbearance during the pandemic. To the extent these frictions are independent of borrower characteristics or mortgage performance, they provide ample quasi-random variation in forbearance provision, which is leveraged to estimate forbearance's causal effects.

2.2.3 Composition of CARES Forbearance Users

Generally, homeowners who opted into CARES forbearance were seeking relief from financial hardship: eventual opters were already over 10 percentage points more likely to be behind on mortgage payments than non-opters, and 12 percentage points more likely to have been in forbearance prior to the introduction of the CARES act. In addition, 1 shows forbearance

[6]From an optimal policy perspective, one could argue that improving household awareness of forbearance options might have enhanced financial outcomes for more borrowers. However, such a counterfactual increase in forbearance uptake could have strained servicer liquidity, with potential aggregate implications. Further research is needed to assess whether expanding the use of forbearance during COVID-19 would have increased systemic risk

opters had weaker credit profiles than the overall sample of federally backed homeowners. Prior to COVID, the eventual forbearance users had credit scores[7] 50 points lower than the overall sample, with 50% higher credit card balances, and 68 p.p. higher revolving credit utilization rates. The forbearance group also contains a higher percentage of FHA/VA loans (50% versus 33% in the overall sample). These facts point to significant selection on borrowers entering forbearance, and financial need being a primary motivator. During COVID, many servicers extended generous forbearance allowances for non-conforming mortgages, which were not covered under the CARES Act.[8] Because the CARES mandate did not apply to non-conforming loans, they are excluded from the sample. Cherry et. al discuss the effects of COVID-era forbearance on jumbo loans, which make up a significant portion of non-conforming loans. For a more detailed description of the composition of COVID-19 mortgage forbearance recipients, see Farrell et al. (2020)

3 Relevant Literature

This paper's novel contributions are to document forbearance's long term effects on the balance sheets of a representative sample of all federally backed homeowners and show how these effects vary by exit plan. This work builds primarily on a growing literature of papers assessing the impact of mortgage and other forbearance programs during COVID. The most closely related paper is Kim et al. (2022), which employs a similar judges IV design to estimate forbearance's causal effect on a selected subset of borrowers, in particular FHA and VA mortgaged homeowners, for a shorter window, between 2020 and 2021. Another important paper in this literature is Cherry et al. (2021), which provides a detailed comparison of COVID-era forbearance policies on mortgages and other loans. In addition, they estimate a causal effect of CARES forbearance around the conforming loan limit. Relatedly, An et al. (2022) shows that minority and low-income households were more likely to use forbearance than their white, higher-income peers, but were more likely to fall behind on payments after forbearance. The authors also show evidence that 40-year loan modifications may have provided vulnerable households with more effective, long term debt relief. Albuquerque and Varadi (2022) and Adelino et al. (2024) document heterogeneous impacts of mortgage forbearance on households of varying income and saving levels using data from the UK and Portugal, respectively. Dinerstein et al. (2023) examines the impact on consumer credit of the student loan moratorium and finds that recipients used the liquidity to increase debt on new loans rather than avoid delinquencies. Gerardi et al. (2022) analyzes the impact of forbearance in tandem with the Federal Reserve's expansionary monetary policy, which lowered mortgage rates and allowed many homeowners to reduce monthly payments through refinancing. Similar to this paper, Gerardi et al. (2022) finds that CARES forbearance was highly effective at reducing financial distress and foreclosure risk in a favorable macroeconomic environment. However, they caution that the results may not port as well to contexts with less favorable housing market conditions, in which borrowers would have had less of a cushion from home equity.

This paper also complements the growing literature comparing the efficacy of mortgage relief policies during COVID with the housing policy response to the Great Recession. The home equity trends documented by Gerardi et al. (2022) are particularly important for comparing policy impacts during COVID and the Great Recession because household debt was exceptionally

[7]Credit scores are based on the VantageScore 4.0 model.

[8]Non-conforming mortgages do not meet the government's underwriting guidelines, such as limits on the original loan volume, debt-to-income ratio, FICO score, down payments, and other requirements.

high during the latter period, and it was notoriously pinned as a cause and catalyst of the Great Recession by Mian and Sufi (2010), and Mian and Sufi (2014). Furthermore, Gerardi et al. (2022)'s findings on the efficacy of forbearance in the context of refinancing due to the low mortgage rate environment complement earlier work on the effects of expansionary monetary policy and quantitative easing on housing markets during the Great Recession, as discussed by Di Maggio et al. (2020). In addition, Gerardi et al. (2022) show that relative to Great Recession area mortgage relief programs– such as the Home Affordable Refinance Program and the Home Affordable Modification Program (HAMP)– where servicer vetting and extensive paperwork processes often led to a year lag between application and relief delivery, the minimal paperwork requirements for CARES forbearance, combined with the universal interest rate reductions led to effective, more immediate support for impacted households.

From a hands-on policy perspective, the COVID-era Director of the Federal Housing Finance Agency, Mark Calabria, commissioned with designing and rolling out mortgage relief policies during COVID affirms Gerardi et al. (2022) in his practitioner's guide detailing his role in designing the mortgage forbearance section of the CARE Act.(Calabria, 2023). In the guide, Calabria attributes the improved performance of CARES forbearance over HAMP and HARP to reducing the paperwork requirements, removing means-tested conditions from the relief, extending the relief duration for up to 12 months (in some cases 18 months) and requiring the relief to be paid back. Calabria also discusses the necessity to balance avoiding "payment shock" for borrowers exiting forbearance, akin to the default risk associated with ARMs rate adjustment during the Great Recession, with the need to protect the balance sheets of the GSEs by ensuring timely repayment.[9] One of the novel contributions of my paper is to illustrate that, by allowing servicers full discretion over workout plans, policymakers may have missed opportunities to protect mortgage stability after forbearance. In practice, this discretion negatively impacted borrowers with balloon payments, which arguably backfired for servicers. In the heterogeneity analysis of this paper, I show that balloon payments were associated with persistently higher delinquency rates relative to forbearance exits with gradual workout plans. This represented an immediate cost to borrowers and servicers, and exposed the GSEs to potential risk in cases of subsequent default.

This paper contributes novel insights to the household finance literature, which examines the sources of financial distress and identifying effective policies for its remediation. Dobbie and Song (2015) and Dobbie et al. (2017) employ a judges IV design to show the long term impacts of filing for chapter 13 bankruptcy. These papers find that bankruptcy boosts 5-year income and credit scores and reduces adverse financial events, mortality, foreclosures and other adverse credit events. This paper adapts their identification strategy to study the causal impact of mortgage forbearance, exploiting quasi-random variation in the generosity of mortgage servicers. Dobbie and Song (2020) provide evidence from a large-scale RCT that interest rate write-downs were significantly more effective than immediate payment reductions at improving financial and labor market outcomes, despite not taking effect for several years. This paper's findings contrast with Dobbie and Song (2020) by showing that the short term liquidity boost provided by COVID had large and persistent positive effects on balance sheet health. Keys et al. (2023) shows that collections and defaults are largely driven by person-specific factors, while bankruptcies are largely place based and explain a significant share of the geographical

[9]Generally, when a GSE-backed mortgage becomes delinquent and is subsequently modified or deferred, it is removed from the pool of collateral used for mortgage-backed securities and placed directly on the balance sheet of Freddie or Fannie. Forbearance avoids this process and the implied risk exposure to the GSE as long as the mortgage continues to perform after payments resume.

variation in financial distress across the country. Finally, Ganong and Noel (2022) decompose the motives of defaulting on a mortgage and show that the vast majority are driven by negative life events (70%), while 24% are driven by a combination of negative life events and negative equity, i.e., underwater mortgages, while the remainder of defaults are driven solely by the latter strategic motive. This finding provides important context for the effects of mortgage forbearance on reducing foreclosure, which should be interpreted through the lens of reducing household's exposure to potentially life-altering adverse events.

Finally, this paper relates to the macroeconomics literature examining optimal countercyclical stabilization policy. Lee and Maghzian (2023) and Auclert et al. (2019) discuss targeted debt relief as countercyclical stabilization policy. Campbell et al. (2021), Guren et al. (2021), and Altunok et al. (2023) discuss the transmission of monetary policy through housing markets with an eye towards optimal mortgage design.

4 Data Description

The data is provided by California Policy Lab's University of California-Consumer Credit Panel (UC-CCP).This dataset contains highly detailed information from the credit reports of a 2% nationally representative sample of American borrowers from 2004Q1 to the most recent quarter (2023Q4). The data is provided at the loan level and comprises detailed payment records and loan characteristic of mortgages, credit cards, home-equity lines of credit, auto loans and leases, student loans, as well as other tradeline categories. Key variables include original loan balances, credit limits, amount past due, scheduled payments, actual payments on selected loans, loan type (Fannie, Freddie, VA/FHA, other), and payment grids delineating the payment histories of each loan at the monthly frequency. I construct mortgage delinquency, distress, and foreclosure indicators based on these payment grids, indexed to the mortgage account balance date. Because several key variables are provided at the quarterly level, I aggregate the monthly mortgage indicators by selecting the maximum within loan and quarter. The data also contains account condition and special comments codes which provide additional details about the loan's performance, such as forbearances, deferrals, modifications, and terminations with specific causes (e.g., refinancing, foreclosure, debt settlements, charge-offs), bankruptcies, and mortgage insurance use. I use the account condition and special comment codes to construct measures of mortgage forbearance, modifications, deferrals, refinancing, and foreclosure starts. Throughout the analysis, I restrict the sample to a balanced panel of $N = 490,710$ primary borrowers with at least one active mortgage in 2020. As noted above, these mortgages are typically held by borrowers with low to moderate income and credit risk and account for approximately 75% of mortgaged homeowners in the data. Excluded from the sample are non-mortgaged homeowners, renters, and borrowers with non-federally backed loans. Below are summary statistics of the sample divided by ever forbearance status.

Table 1 summarizes key variables for all borrowers in the balanced panel. The left panel shows summary statistics for the full sample, including forbearance users, while the right panel Table 1 compares the credit profiles of borrowers in the overall sample with those who entered forbearance following the introduction of the CARES Act, who represent less than 9% of the sample. Across most dimensions, borrowers in the forbearance (treatment) group resemble those in the broader sample. However, a notable exception is that forbearance users were considerably more likely to have FHA or VA mortgages (52% of forbearance users versus 31% of the overall sample). Although the UC-CCP data lacks direct information on income and savings, the distinctive features of FHA and VA loans—such as lower down payment and

Table 1: Summary Statistics of All UC-CCP Federally Backed Homeowners

	Full Sample			Forbearance Users		
	Mean	Std Dev	Median	Mean	Std Dev	Median
Pre-COVID						
Credit score	760.07	78.18	785.00	770.91	72.38	796.00
Credit card balance	5,115.98	8,975.83	1,569.00	5,086.11	9,112.79	1,475.00
RC utilization rate	0.26	0.75	0.10	0.24	0.79	0.08
Credit limit (all accounts)	22,736.61	21,634.73	16,750.00	24,699.62	22,964.56	18,524.00
Original mortgage bal.	219,138.30	120,345.94	194,660.00	223,750.28	125,948.09	196,886.00
Post-COVID						
Credit score	688.45	94.20	693.00	702.67	91.85	707.00
Credit card balance	7,013.30	10,800.56	3,264.00	7,086.45	10,775.47	3,151.00
RC utilization rate	0.46	0.88	0.37	0.42	1.10	0.30
Credit limit (all accounts)	19,298.15	20,876.76	12,501.00	21,482.07	22,661.10	14,450.00
Original mortgage bal.	223,750.28	125,948.09	196,886.00	232,963.83	122,078.27	208,000.00
FHA/VA status	0.31			0.52		
Female	0.47			0.46		
Self employed	0.00			0.00		

credit score requirements, offset by higher interest rates—suggest that these borrowers are more likely to be lower-income, liquidity-constrained, and often first-time homeowners.(Cherry et al., 2021)

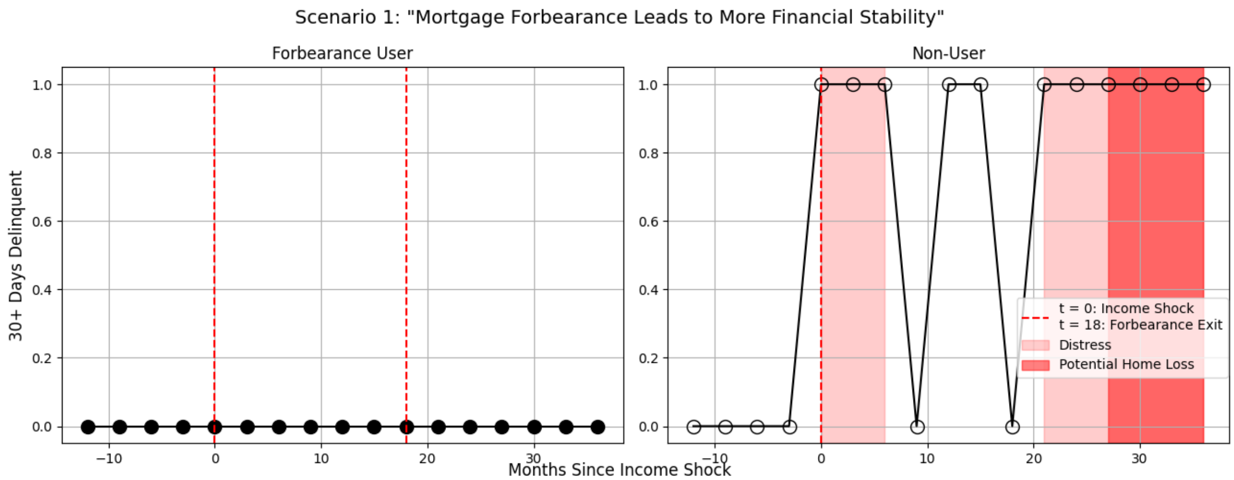
5 Research Design

I estimate the causal effect of forbearance on financial stability outcomes via two-stage least squares, with plausibly exogenous differences in servicer propensity as an instrument for forbearance. To motivate this design, consider an ideal experiment in which forbearance is randomly provided to households experiencing negative income shocks at t = 0 and lasts for 18 months. In the ideal experiment, the randomization of treatment permits a causal interpretation to the resulting differences in outcomes after forbearance ends at t=18. The diagram below illustrates potential outcomes for non-zero treatment effects at long horizons (beyond t = 18), using mortgage delinquency as the primary response variable:

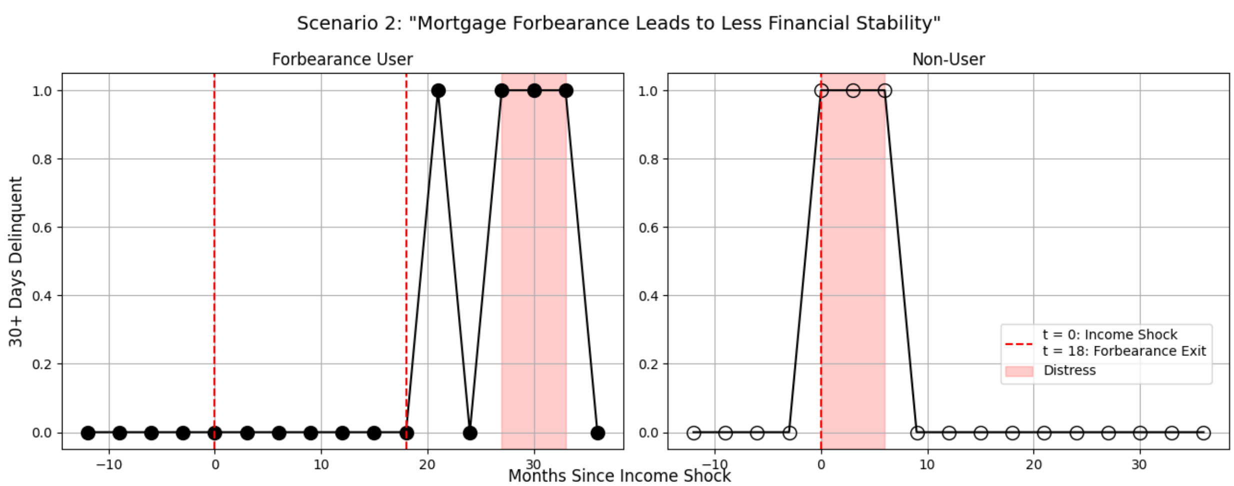
The diagram shows plausible cases where the true long-term causal effect of forbearance on financial stability could be positive or negative.[10] Figure 1a illustrates the case where forbearance prevents mortgage delinquency beyond relief expiry (t =18), compared to a control group, where both the treatment and control groups face negative income shocks at t=0. In this scenario, the control group faces higher risk of sustained delinquency and potential foreclosure, driving the non-zero treatment effect at long horizons. Figure 1b illustrates how forbearance could postpone delinquency until after relief expiry, followed by a rebound effect. In this scenario, the control group benefits from alternative strategies to temper the effects of the initial shock. Non-zero treatment effects may result from substitution between forbearance

[10]It is also plausible to obtain a null treatment effect.

Figure 1: Competing Hypotheses of the Causal Effects of Forbearance



(a) This panel illustrates the case where forbearance prevents mortgage delinquency beyond relief expiry (t =18), compared to a control group, where both the treatment and control groups face negative income shocks at t=0. In this scenario, the control group faces higher risk of sustained delinquency and potential foreclosure, driving the non-zero treatment effect at long horizons.



(b) This panel illustrates how forbearance could postpone delinquency until after relief expiry, followed by a rebound effect. In this scenario, the control group benefits from alternative strategies to temper the effects of the initial shock. Non-zero treatment effects may result from substitution between forbearance and other, more effective loss mitigation programs, behavioral responses to treatment (or non-treatment), or accumulating debt during forbearance.

and other, more effective loss mitigation programs, behavioral responses to treatment (or non-treatment), or accumulating debt during forbearance.

In practice, forbearance provision was not random during the pandemic. Instead, the literature indicates that borrowers with lower income, greater financial risk, and minorities were significantly more likely to enter forbearance under the CARES Act.(An et al., 2022; Cherry et al., 2021; Farrell et al., 2020). To the extent that borrower selection into treatment is unobserved and correlated with outcomes, a simple OLS estimate of the treatment effect would be biased. For example, the estimated effect of forbearance on reducing mortgage delinquency could be overstated if borrowers with higher financial literacy are less likely to be delinquent in general but more proactive in using forbearance during COVID. Conversely, the efficacy of forbearance could be underestimated if unobserved income shocks are correlated with both higher forbearance rates and subsequent delinquencies. To distinguish between the competing hypotheses in Figure 1, I use an instrumental variables approach that leverages plausibly exogenous variation in mortgage servicers' propensity to provide forbearance during the pandemic. The identification strategy posits that after controlling for borrower characteristics and local economic conditions that drive forbearance decisions, the residual variation in mortgage servicers' willingness to provide forbearance is as-good-as random. Formally, I estimate a two-stage least squares (2SLS) with servicer propensity as an instrument. The second stage estimating equation is a local projection of $EverForbearance_i$ on outcomes at period τ:

$$y_{i,\tau} = \alpha_{z,\tau} + X_i'\Gamma_\tau + \beta_\tau Ever\widehat{Forbear}ance_i + \epsilon_{i,\tau} \tag{1}$$

β_τ represents the causal effect of forbearance on the outcome at the τth horizon. τ representing the month of the end of each quarter after the CARES Act was introduced in March 2020 ($\tau = 0$). Each β_τ is estimated using a cross-sectional regression that incorporates outcome data at time τ, borrower characteristics, X_i, forbearance status, $EverForbearance_i$, and zip code fixed effects, $\alpha_{z,\tau}$.

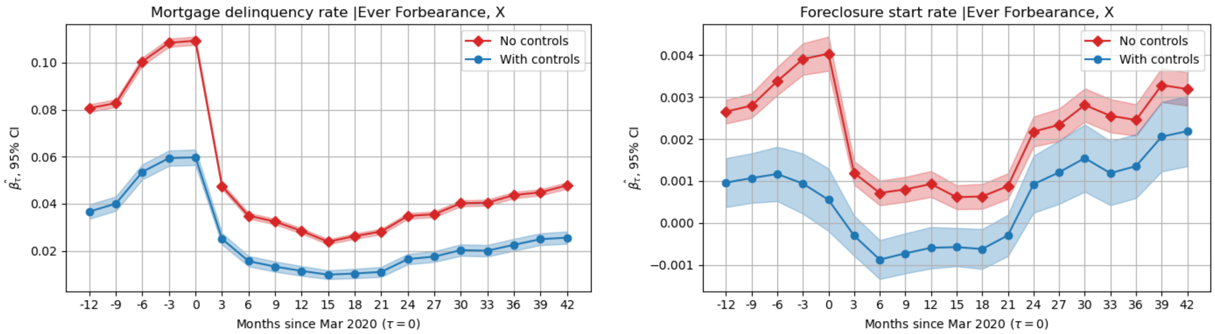
X_i includes time-invariant borrower i characteristics, such as credit score, mortgage type (GSE, FHA/VA), original loan volume, gender, self-employed status, revolving credit utilization, and credit card limits. $EverForbearance_i$ is a time-invariant indicator for whether an individual ever received forbearance after the CARES Act was introduced.

The corresponding first-stage estimating equation is:

$$EverForbearance_i = \delta_z + X_i'\Delta + \gamma\hat{\sigma_s} + \eta_i \tag{2}$$

$\hat{\sigma_s}$ is the estimate of the systematic component of servicer behavior, which describes the servicer's marginal impact on the probability an eligible borrower ever receives forbearance under the CARES mandate. σ_s contains the conditionally exogenous factors influencing servicer forbearance provision, such as whether a servicer is a shadow bank or a federally regulated depository institution, financial constraints, risk tolerance, degree of capitalization, organizational form, size, as well as COVID-specific information and logistical frictions around delivering forbearance as noted in the introduction. Cherry et al. (2021), Kim et al. (2022) and Aiello (2022) document these factors as part of the key role mortgage servicers played in determining distressed mortgage outcomes during COVID and the Great Recession.

Figure 2: Mortgage Stability Outcomes (OLS)



(a) Mortgage Delinquency Rate

(b) Mortgage Foreclosure Start Rate

Panel 2a shows the OLS estimated coefficients on mortgage delinquency from ever receiving CARES forbearance with and without additional controls. Panel 2b shows the differential foreclosure rate. In both panels, the blue lines represents the estimated rate after controlling for borrower and mortgage characteristics. The red lines represent the average rates without including controls. Including controls reduces but does not eliminate the presence of pretrends.

6 Instrument Description

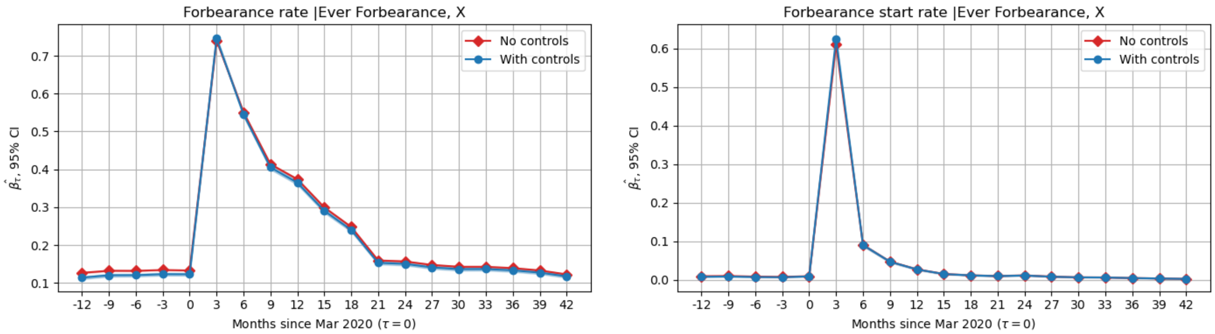
As described above, confounders such as income and financial literacy may bias the OLS estimates of the causal effect of forbearance. I provide empirical evidence of borrower selection and discuss the proposed instrument's relevance and plausible exogeneity.

6.1 OLS Findings and Caveats

The estimates presented below provide evidence of borrower selection due to unobserved risk factors, which may include financial literacy, negative income shocks, myopia, liquidity constraints, or servicer-side factors. In Figure 2 and 3, I estimate the local projection coefficients β_τ on ever receiving forbearance under the CARES Act using cross-sectional OLS regressions. The red line plots simple average differential rates, while the blue line plots OLS estimated differential rates, which are adjusted by orthoganlizing to a rich set of borrower and mortgage characteristics and zip code fixed effects to account for observed differences across borrowers. In Figure 2, the inclusion of controls reduces but does not eliminate the pre-trend estimates of the difference in mortgage delinquency and foreclosure start rates. The endurance or pre-trends after including controls is not entirely surprising given the close resemblance of borrowers in the treatment and overall sample as shown in Table 1. It is worth noting that while the borrowers in the treatment group are largely similar, a key distinguishing feature is that they are significantly more likely to hold a FHA or VA mortgage instead of a GSE-backed loan. This difference may shed light on specific unobserved confounders at play, such as liquidity constraints, lower income, or lower wealth.

The contrast between the red and blue lines in Figure 2 indicates that borrower selection is a significant factor and can be partially addressed by controlling for observed characteristics from the credit reports. For comparison, I include the β_τ estimates of the timing of forbearance below. The red and blue lines largely overlap, demonstrating that the timing of forbearance is

Figure 3: Timing of Forbearance (OLS)



(a) Rate of Participants In Forbearance

(b) Rate of Participants Starting Forbearance

Panel 3a shows the rate of CARES participants in forbearance at each point in time. For example, the rate of .75 at $\tau = 3$ indicates that among borrowers who ever used forbearance after CARES was introduced, 75% were in forbearance during 2020Q2. Panel 3b shows the rate of forbearance starts among CARES users at each horizon. The close overlap between the red and blue lines indicates that the timing of mortgage forbearance does not vary significantly with borrower or mortgage characteristics.

largely independent of regional, borrower and mortgage risk factors. Similar to the mortgage stability plots, the timing plots below reveal statistically significant pretrends. These may be driven by unobserved borrower factors, or more likely servicer-side factors, which are not included as controls due to limitations in the data.

Using the 2SLS design described above, I construct the instrument using servicer fixed effects σ_s from the following cross-sectional regression:

$$EverForbearance_i = \phi_z + X_i'\Gamma + \sigma_s + \epsilon_i \tag{3}$$

where ϕ_z is a zip code fixed effect, X_i is a vector of borrower controls and σ_s is the systematic component of the servicer's impact on forbearance probability. The figures below describe the distribution and relevance of the servicer generosity instrument.

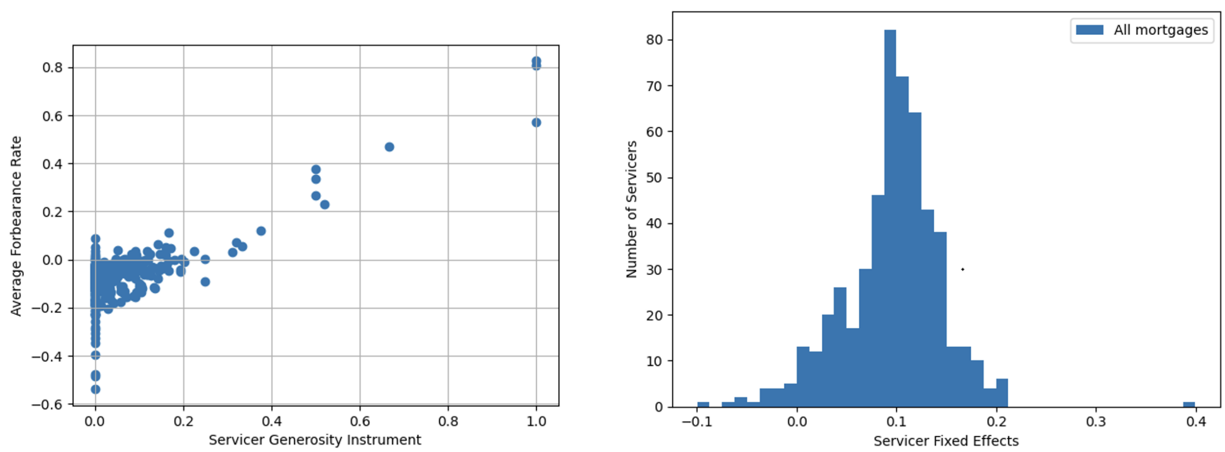
Figure 4a shows the within-servicer rate of forbearance by estimated forbearance generosity. The scatter plot shows the instrument is highly relevant to forbearance decisions, on average. Figure 4b plots the histogram of servicer generosity levels estimated from equation 3, revealing significant variation in servicer's willingness to provide forbearance during COVID, after orthoganlizing to borrower and mortgage characteristics and zip codes. The variation is interpreted as being driven by servicer logistical and informational frictions as noted in the Fannie Mae surveys in the introduction, or servicer liquidity constraints, organizational form, and regulatory pressures as in Kim et al. (2022) and Cherry et al. (2021). In the table below, I present evidence that the instrument has near-zero correlation with an extended set of borrower and mortgage characteristics. The test of randomness is constructed using the following regression:

$$\hat{\sigma_i^s} = X_i'\Omega + \mu_z + \varepsilon_i \tag{4}$$

Figure 4: Servicer Generosity Instrument

(a) Servicer Generosity and Forbearance Rate

(b) Variation in Servicer Generosity



Panel 4a shows the within-servicer rate of forbearance by estimated forbearance generosity. Panel 4b plots the histogram of servicer generosity levels estimated from equation 3

Coefficient estimates near zero suggest the instrument is uncorrelated with borrower and mortgage characteristics, and instead represents plausibly exogenous variation driven by servicer factors.

7 Results

The figures below represent the estimated causal effects of forbearance provision on key outcome variables using the 2SLS model specified in equations 1 and 2. 95% confidence bands are also shown, with all standard errors clustered at the zip code level.

Figure 5 shows the timing of forbearance among CARES participants, using the IV specification in equation 1. Here, $\tau = 0$ corresponds to 2020Q1, when the CARES Act was first introduced. During COVID, forbearance uptake was highly concentrated in the 18-month period between $\tau = 3$ (2020Q2) and $\tau = 21$ (2021Q4). The β_τ for $\tau \geq 21$ represent post-forbearance outcomes, as the majority of participants had exited forbearance by this time. The $\tau = 21$ threshold is particularly relevant for borrowers who exited early, given the paper's focus on the longer term impacts of forbearance. Below, I present the instrumented causal effects of CARES forbearance on mortgage stability, additional loss mitigation techniques, and revolving credit performance.

Figure 6 presents the impact of forbearance on two key measures of mortgage stability: serious delinquency and foreclosure. Panel 6a focuses on delinquencies of 60 or more days for each borrower's largest active mortgage on file. The depict a steep reduction (4.5 percentage points) in serious delinquency after $\tau = 27$, which persists for the remainder of the sample. These findings point to a strong, positive impact of forbearance on long-term mortgage payment stability. Furthermore, considering that a single mortgage delinquency can lower a borrower's credit score by 50 to 100 points and incur late fees of 3-6% of the past due payment, and that

Table 2: Test of Randomness

	(1)
FHA/VA mortgage	-0.009*** (0.000)
Number of accounts past due	0.003** (0.000)
RC balance 2018	0.003*** (0.000)
Borrower age	0.001*** (0.000)
Female	0.000 (0.001)
Self employed	-0.001 (0.001)
Mortgage age	-0.001* (0.001)
Original mort. balance	-0.016*** (0.000)
Vantage score	0.000*** (0.000)
Credit limit	-0.000*** (0.000)
Average Scheduled Mort. Pymnt	-0.000*** (0.000)
RC balance in Mar20	0.000*** (0.000)
R-squared	0.193
S.E. type	by: zip_cd
Observations	385284

$$\hat{\sigma}_i^s = X_i'\Omega + \mu_z + \varepsilon_i$$

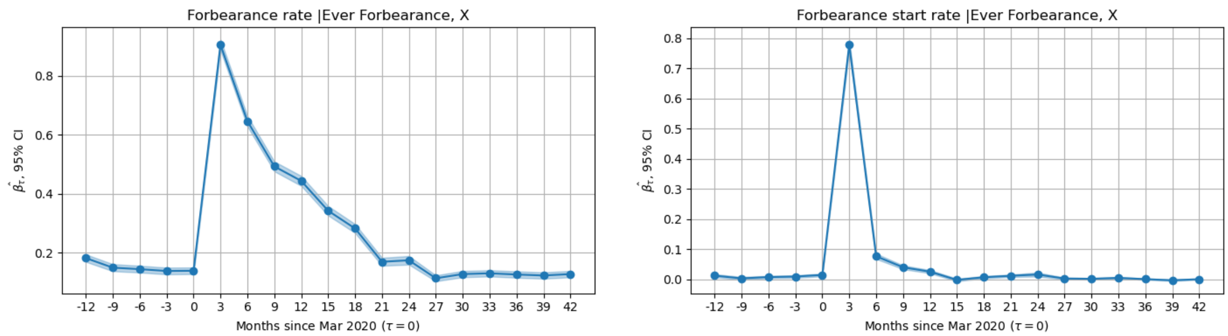
Table 2 shows coefficient estimates from a cross-sectional regression of the servicer fixed effects on an extended set of borrower and mortgage characteristics. Although several point estimates are statistically significant, the near-zero values indicate a very low correlation between the instrument and demand-side factors.

repeated delinquencies (90+ days) often lead to foreclosures, the CARES forbearance provision significantly improved long-term balance sheet health by preventing mortgage delinquency. Panel 6b focuses on the impact of forbearance on foreclosure risk. The results indicate a modest but statistically significant effect on reducing foreclosure rates at long horizons. Taken together, these findings demonstrate that the temporary payment pause during COVID had a significant, lasting impact on improving household financial health through stabilizing mortgages. These results are consistent with the model in which forbearance enhances financial stability by providing timely relief to borrowers.

Under the CARES Act, the levels of relief from forbearance varied significantly, as mortgage servicers retained discretion over the workout plans used to bring mortgages current after the forbearance period. These plans ranged in leniency, from balloon payments (also known as re-instatements), where borrowers were expected to settle the entire deferred amount immediately after forbearance, to more gradual repayment plans, allowing homeowners an extended period to resolve their outstanding payments. Additionally, servicers were encouraged to screen forbearance candidates before granting relief, determining whether an alternative loss mitigation solution might be more effective.

The interaction between forbearance and other loss mitigation strategies highlights best practices for reducing financial distress during recessions through mortgage relief programs.

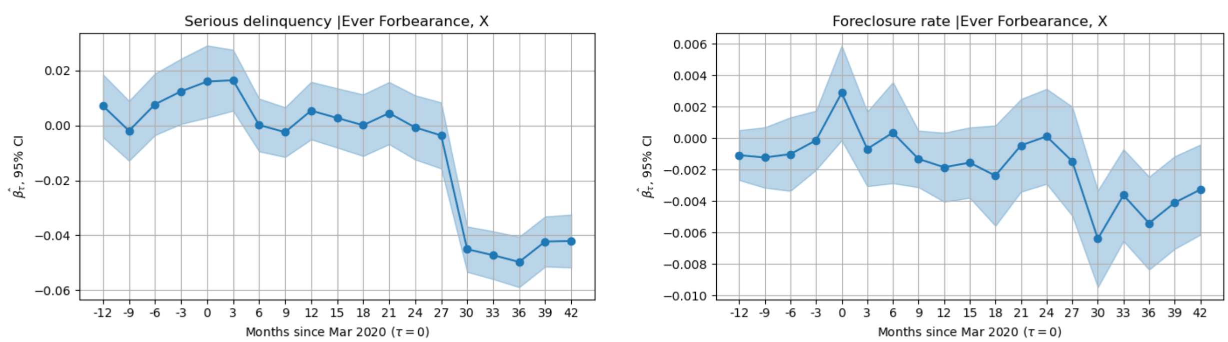
Figure 5: Timing of Mortgage Forbearance



(a) Rate of Participants in Forbearance (b) Rate of Participants Starting Forbearance

Panel 5a shows the proportion of participants using the program at each point time. Panel 5b shows the start rate among forbearance users. Roughly 80% of participants start forbearance at $\tau = 3$ (2020Q2), with the vast majority of borrowers exiting by $\tau = 21$ (2021Q4). After $\tau = 21$, the forbearance rate among CARES users returns to the pre-pandemic level, around 15%.

Figure 6: Mortgage Stability Effects



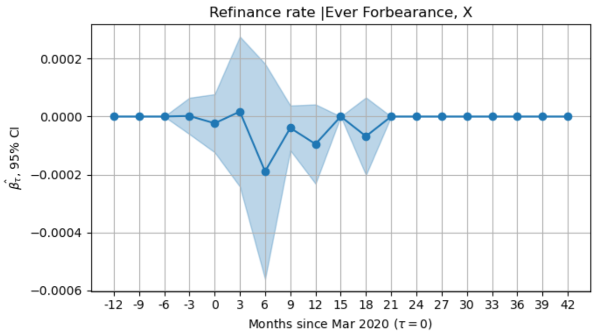
(a) Mortgage Delinquency Rate (b) Mortgage Foreclosure Rate

Panel 6a shows the instrumented causal impact of forbearance on mortgage delinquency risk using the largest active mortgage on file for each borrower. Delinquency is defined as any payment more than 30 days past due. Panel 6b shows the instrumented causal impact on the foreclosure rate. Both graphs show statistically insignificant pretrends, indicating a negligible correlation between the servicer generosity instrument and mortgage stability outcomes prior to COVID.[11]

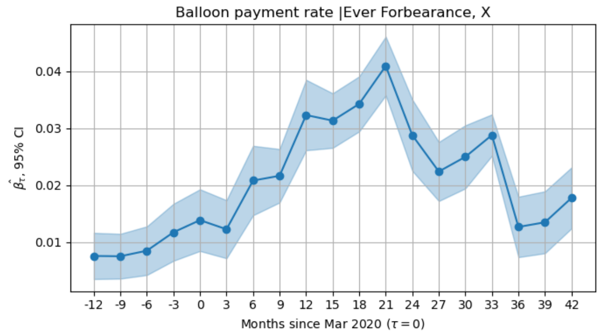
Below, I present findings on how forbearance influenced mortgage servicers' provision of loss mitigation alternatives.

Figure 7 the effects of forbearance provision on the use of loss mitigation alternatives, in particular, refinancing in Panel 7a and balloon payments (reinstatements) in Panel 7b. Forbearance leads to a noteworthy reduction in refinancing rates between $\tau = 6$ and $\tau = 21$, with

Figure 7: Loss Mitigation Strategies



(a) Mortgage Refinance Rate



(b) Mortgage Balloon Payment Rate

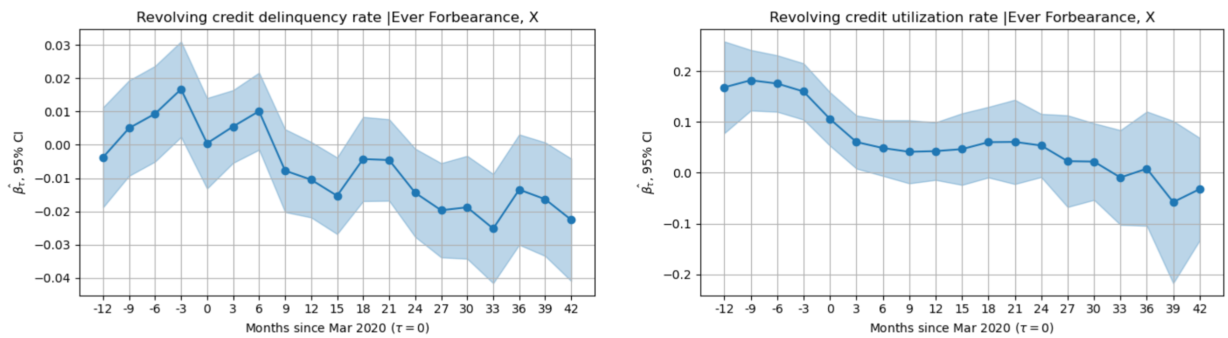
Panel 7a shows the instrumented effect of CARES forbearance on refinance rates, where sufficient variation is available to estimate β_τ. Panel 7b shows the instrumented effect on balloon payment rates.

no significant impact at longer horizons. The substitution between forbearance and refinancing at earlier horizons is mechanical while borrowers were in forbearance, and it highlights a potential unintended consequence of the relief: by opting into forbearance, borrowers may have missed the opportunity to refinance into a lower interest rate. Despite these potential foregone savings, the forbearance recipients' gains in terms of mortgage stability suggest that the timeliness of forbearance relief may have been particularly influential, despite its temporary structure. Panel 7b shows how forbearance provision affected the issuance of balloon payments around COVID. Interestingly, servicers with more generous forbearance provision during COVID were slightly more likely to issue balloon payments prior to the pandemic. Balloon payment rates peak at $\tau = 21$, when the last cohort of participants exited forbearance. While the spike in balloon payments could have potentially caused a "rebound effect" in terms of mortgage default, the empirical evidence on mortgage stability outcomes in Figure 6 demonstrate that forbearance recipients retained their ability to meet mortgage payment obligations after leaving forbearance. These findings provide additional evidence of the forbearance program's efficacy despite the potential lost refinancing opportunities and financial stability risk imposed by balloon payments.

The empirical findings demonstrate the effectiveness of forbearance under the CARES Act in stabilizing mortgage performance during the COVID-19 recession. A natural follow-up question is whether recipients genuinely needed forbearance or if they used the program as an opportunity to increase short-term liquidity and consumption. However, as illustrated below, the provision of forbearance actually resulted in a reduction in revolving credit utilization rates. This suggests that recipients did not channel the additional liquidity into increased spending, indicating that opportunistic use of forbearance had a limited impact on the mortgage stability outcomes described above.

Panel 8a shows that forbearance had a modest, statistically significant impact on reducing credit card delinquencies between $\tau = 24$ and $\tau = 33$. Panel 8b indicates that revolving credit utilization rates fell in response to forbearance provision. Taken together, the revolving credit response to forbearance provides evidence of positive spillovers in terms of reducing credit

Figure 8: Revolving Credit Stability Effects



(a) Revolving Credit Delinquency Rate

(b) Revolving Credit Utilization Rate

Panel 8a represents the instrumented effect of CARES forbearance on credit card delinquency rates. The credit card delinquency rate is defined as available to estimate β_τ. Panel 8b shows the instrumented causal impact on balloon payment rates.

card delinquencies, without prompting additional spending in response to the liquidity boost.

8 Heterogeneity Analysis

In this section, I analyze the heterogeneous impacts of mortgage forbearance based on exit plans and forbearance duration. As highlighted by Calabria (2023), policymakers faced the challenge of designing forbearance policies that minimized "payment shock" for borrowers resuming payments, while also protecting servicers from liquidity risk and GSEs from credit risk. Key dimensions of the policy included the structure of repayment and the length of forbearance allowed under the policy. The CARES forbearance provision gave servicers full discretion over forbearance workout plans, with the knowledge that servicers were incentivized to provide borrowers with sustainable post-forbearance plans. Additionally, the CARES Act permitted up to 12 months of forbearance, prioritizing borrower relief to enhance the program's stimulative effects, though this also increased the risk of "payment shock" for borrowers who opted for extended forbearance periods.

8.1 Event Study Design

To study the heterogeneous effects of forbearance by exit plan and duration, I restrict the sample to borrowers who used forbearance after CARES was introduced with a known exit data recorded in the data. Then, I stack the observations to $k = 0$ at the monthly period when each borrower exits forbearance. I further restrict the sample to a balanced panel of borrowers who are observed between $k = -18$ and $k = 36$. The specification is:

$$y_{i,t} = \alpha_t + X_{i,t}\gamma + \sum_{k=-18,k\neq-3}^{36} \beta_k D_{i,t}^k \times Feature_i + \varepsilon_{i,t} \tag{5}$$

β_k captures the differential impact of having a particular forbearance policy feature (balloon payment or long forbearance) at at period k relative to the control group. $X_{i,t}$ includes credit scores, number of accounts past due, and credit card balances prior to COVID. α_t is a calendar time fixed effect.

8.2 Balloon Payments vs Other Exit Plans

Because forbearance programs outside the COVID context typically require lump-sum balloon payments, many borrowers and servicers mistakenly assumed this would apply to CARES Act forbearance provisions as well, which in some cases deterred borrowers from opting in (Consumer Financial Protection Bureau, 2021).[12] When considering the trade-off between protecting household financial stability and managing risk to servicers, balloon payments help ensure prompt repayment to servicers but can increase the risk of delinquency, as these payments generally must be made in full. To examine whether balloon payments caused borrowers to miss mortgage payments at a higher rate, I use the the event study specification in equation 5, with $Feature_i = 1$ if the borrower ever encounters a balloon payment after exiting forbearance.

Below I present the summary statistics of balloon forbearance exiters, compared with the control group of non-balloon exiters. The sample is restricted to a balanced panel of borrowers observed between $k = -18$ and $k = 36$.

[12] In practice, however, UC-CCP data shows that only 9% of CARES forbearance users encountered balloon payments after exiting forbearance.

Table 3: Summary Statistics by Forbearance Exit Type

	Non-Balloon Exit			Balloon Exit		
	Mean	Std Dev	Median	Mean	Std Dev	Median
Pre-COVID						
Credit score	688.45	94.20	693.00	702.67	91.84	707.00
Credit card balance	7,013.44	10,800.67	3,264.00	7,086.45	10,775.47	3,151.00
Revolving credit utilization rate	0.46	0.88	0.37	0.42	1.10	0.30
Credit limit (all accounts)	19,298.52	20,876.88	12,501.00	21,482.45	22,661.07	14,450.00
Post-COVID						
Credit score	702.67	91.84	707.00	714.09	92.60	722.00
Credit card balance	7,086.45	10,775.47	3,151.00	7,777.96	11,515.22	3,491.00
Revolving credit utilization rate	0.42	1.10	0.30	0.57	0.28	0.39
Credit limit (all accounts)	21,482.45	22,661.07	14,450.00	23,282.08	22,993.77	16,680.00
Original mortgage bal.	223,732.17	125,924.18	196,886.00	232,963.83	122,078.27	208,000.00
FHA/VA status	0.52			0.52		
Female	0.46			0.46		
Self employed	0.00			0.01		

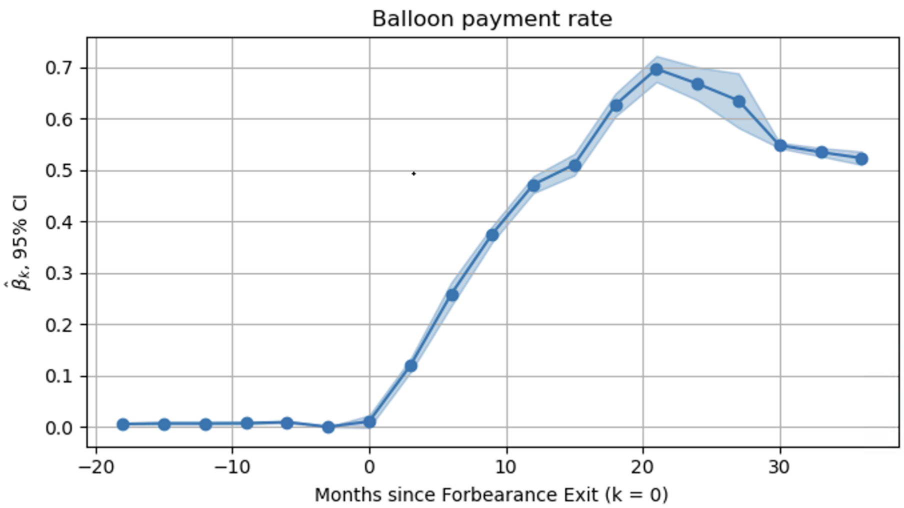
The summary statistics table indicates that the credit profiles of forbearance users with and without balloon payments were broadly similar. However, unobserved factors such as income, savings, and home equity could still influence a servicer's decision to require a balloon payment and affect subsequent borrower outcomes. Consequently, the event study findings should be interpreted as descriptive rather than causal evidence.

Reporting lags likely impact the observed timing of balloon payments in the credit bureau data. According to Figure 9, among balloon exiters, 10% were faced with the lump-sum payments immediately after exiting forbearance, with the rate gradually increasing to 70% within a year and a half of forbearance exit, then gradually falling and tapering off to around 50%. However, policy guidance from the FHFA and other agencies indicates the majority of these balloon payments likely hit immediately as borrowers exited forbearance, and the observed delays are due to reporting lags in the credit bureau data.

Figure 10a plots the β_k estimates from the event study in equation 5, where $Feature_i = 1$ if a borrower ever faces a balloon payment after exiting forbearance. The control group includes borrowers who exited forbearance with workout plans that never featured balloon payments. Upon exiting forbearance, borrowers assigned balloon payments were persistently 2 percentage points more likely to be delinquent than non-balloon exiters. In Figure 10b, balloon exiters were temporarily .1 percentage point less likely to face foreclosure but the effect becomes statistically insignificant at the one year mark. Overall, the fact that balloon payments were associated with persistently higher delinquency risk relative to gradual workout plans indicates that future forbearance programs may improve by restricting the use of balloon payments to a smaller subset with a high likelihood of timely repayment.

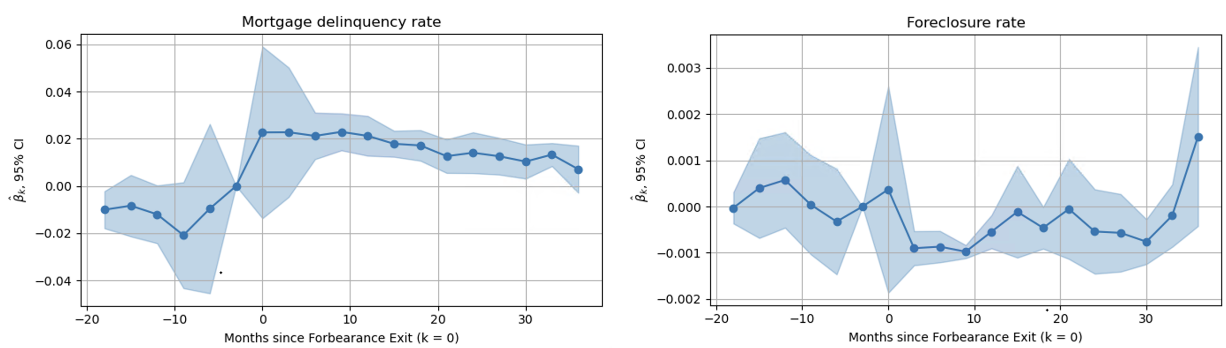
Given that balloon payments increase the risk of delinquency, the naturally ensuing questions would be whether balloon payments correspondingly raise mortgage debts (amounts past due), and whether the risk spills over to revolving credit accounts. In 11, I show a modest,

Figure 9: Timing of Balloon Payments



The graph depicts the differential rate of balloon payments among CARES participants with exit plans featuring balloon payments at some point after forbearance. Each dot point is a β_k estimated from equation 5. The observed delays in balloon payment rates after forbearance exit are most likely driven by reporting lags.

Figure 10: Balloon Forbearance Exit and Mortgage Stability



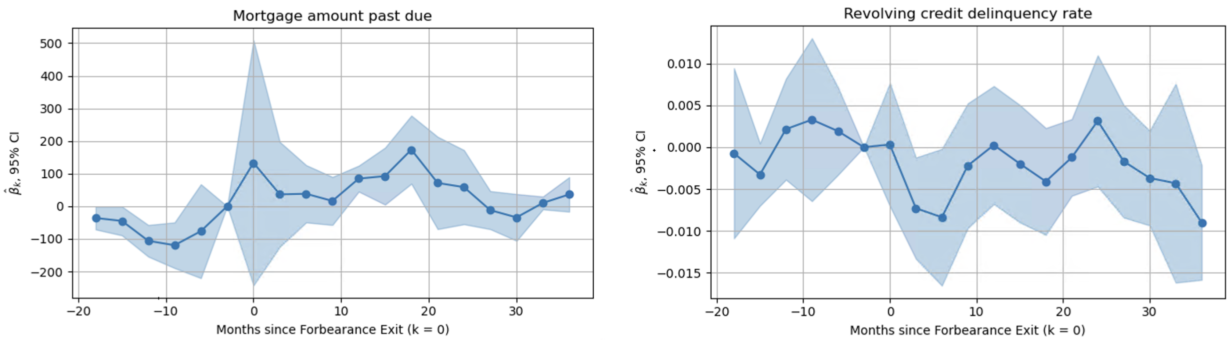
(a) Delinquency Rate with Balloon Exit

(b) Foreclosure Rate with Balloon Exit

Panel 10a plots the differential delinquency rate among balloon exiters. 10b shows the differential foreclosure rate. 95% confidence intervals are also shown, using heteroskedasticity-robust standard errors.

approximately \$100 increase in mortgage amounts past due for balloon exiters, although the effect is statistically indistinguishable from zero for most periods. In addition, I find no evidence of risk spillovers to revolving credit stability. These findings suggest that facing a balloon payment after forbearance did not have an economically or statistically significant impact on debt levels relative to non-balloon exiters.

Figure 11: Balloon Forbearance Exit and Debt Levels



(a) Mortgage Amount Past Due with Balloon Exit (b) RC Delinquency Rate with Balloon Exit

Panel 11a shows the differential mortgage amount past due among balloon exiters. 11b shows the revolving credit delinquency rate. 95% confidence intervals are also shown, using heteroskedasticity-robust standard errors.

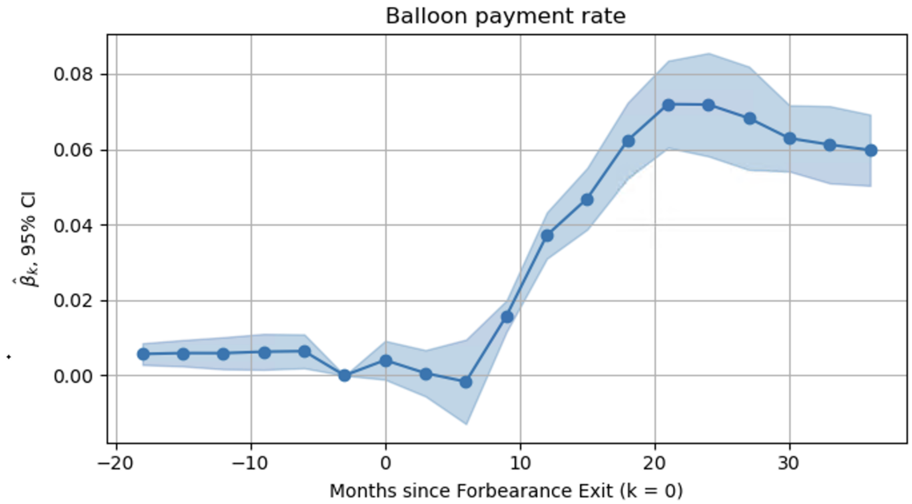
8.3 Long vs. Short Forbearance Duration Effects

The duration and implied volume of relief (as well as repayment obligation) varies significantly across borrowers, with potential implications for borrowers and servicers. In the data, roughly 8% of CARES forbearance participants stayed in forbearance for more than six months. Recall from the competing hypotheses diagram in Figure 1 that in principle, forbearance can potentially lead to financial instability if borrowers accumulate unsustainable debt, form spending habits that do not take into account mortgage habits, or develop a moral hazard. These risks are especially pertinent the longer forbearance goes on. Nevertheless, longer forbearance periods can also be extremely beneficial to households that require these funds to meet housing payment obligations for an extended period, but retain their ability to repay missed obligations once payments resume. To estimate the direction of the effect of having long forbearance versus short forbearance, I employ the event study specification in equation 5. In this analysis, $Feature_i$ is set to 1 if a participant has forbearance for six months or more, and 0 otherwise. As with the balloon payment analysis, unobserved factors such as employment status could influence both the choice of forbearance duration and subsequent financial outcomes, meaning that the findings do not have a purely causal interpretation. Nevertheless, the descriptive results remain relevant for policymakers.

In Figure 12, I plot the β_k estimates from the regression in equation 5, where the covariate of interest is an indicator for whether forbearance lasted for more than 6 months, interacted with event time dummies. The figure indicates that long forbearance users were between 2 and 8 percentage points more likely to face a balloon payment. However, the exact timing of the balloon payments is unclear due to reporting lags in the credit bureau data. The fact that long forbearance users were significantly more likely to be assigned a balloon payments may reflect servicers' eagerness to recoup liquidity after granting extended periods of relief.

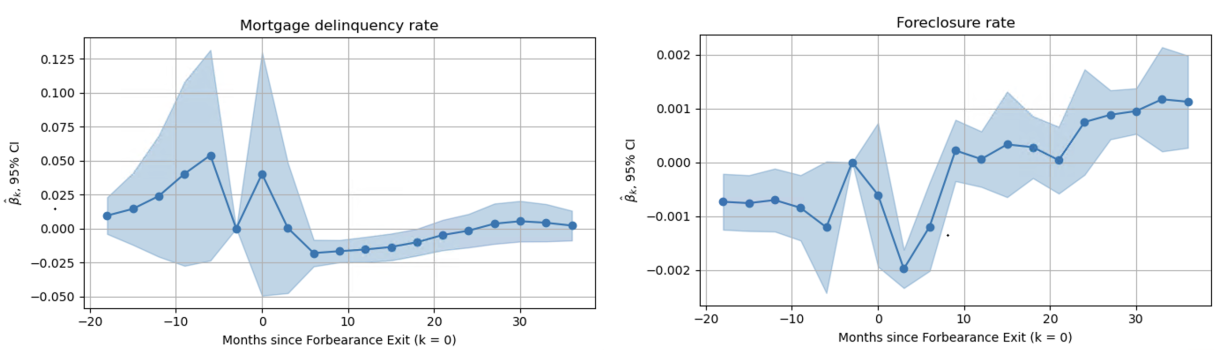
Figure 13 illustrates the differential impact of opting into forbearance for six or more months on mortgage delinquency and foreclosure rates. The results show a temporary reduction in mortgage delinquency rates, followed by a sustained period of no significant effect. The impact

Figure 12: Long Forbearance and Balloon Payments



This figure shows the differential balloon payment rate among forbearance users who stayed in forbearance for greater than six months, relative to those who stayed for fewer than six months. 95% confidence intervals are also shown, with heteroskedasticity-robust standard errors. Long forbearance users were between 2 and 8 percentage points more likely to face balloon payments upon exiting forbearance.

Figure 13: Long Forbearance and Mortgage Stability



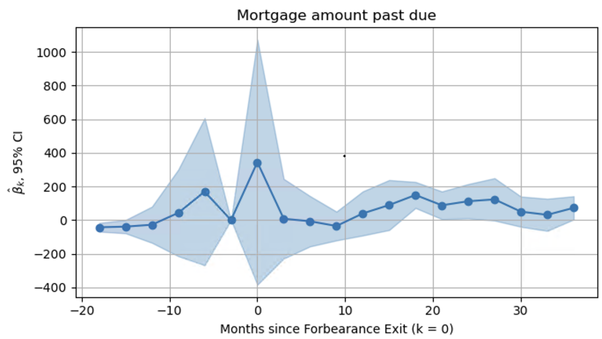
(a) Delinquency Rate with Long Forbearance

(b) Foreclosure Rate with Long Forbearance

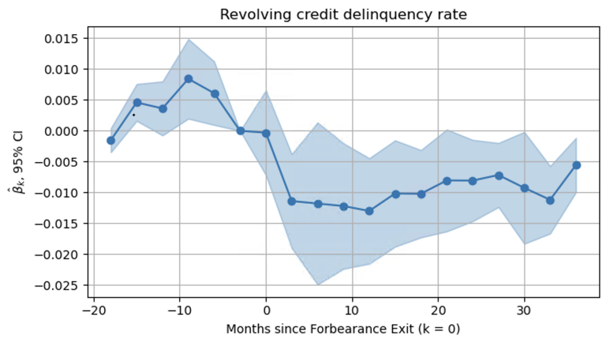
Panel 13a event study results for the differential mortgage rate among long forbearance users. 13b shows the differential foreclosure rate. 95% confidence intervals are also shown, with heteroskedasticity-robust standard errors.

on foreclosure rates is more pronounced, with a substantial reduction of over 1 percentage point one quarter after exiting forbearance, followed by a significant rebound effect that begins to emerge six months after forbearance exit. In terms of distinguishing between the competing hypotheses illustrated in Figure 1, the foreclosure findings support the interpretation that forbearance can lead to instability, especially for borrowers facing persistent financial difficulties, while the mortgage delinquency results lend themselves to the interpretation of a null treatment effect of longer forbearance.

Figure 14: Long Forbearance and Debt Levels



(a) Mortgage Amount Past Due



(b) Revolving Credit Delinquency Rate

Panel 14a event study results for the differential mortgage amount past due among balloon exiters. 14b shows the revolving credit delinquency rate. 95% confidence intervals are also shown, with heteroskedasticity-robust standard errors.

Given the lack of mortgage delinquency risk associated with longer forbearance duration, it is not surprising that dollar amounts past due on the mortgages are similarly unaffected by the length of forbearance, as shown in Figure 13a. However, it is worth noting that longer forbearance duration is associated with a large, persistent drop in revolving credit delinquency rates, which suggests that households use the liquidity from extended relief periods to stabilize their balance sheets. This is consistent with the main finding presented in the results section, in which forbearance users reduce their revolving credit utilization rates as a result of the relief.

Overall, the long forbearance analysis indicates that participants who opted in for extended duration had temporarily lower foreclosure rates, although rebound effects begin to appear after the 6 month mark. Longer forbearance periods were associated with a persistent 1.5pp reduction in revolving credit delinquency rates, with limited or no effect on mortgage delinquency rates or amounts past due. Taken together, the forbearance duration analysis suggests that these borrowers benefited from channeling the liquidity to avoid credit card delinquencies; however, those with entrenched financial hardship remain exposed to foreclosure risk, particularly at longer horizons.

9 Conclusion

This paper employs a judges IV design to estimate the causal impact of COVID-era mortgage forbearance on the financial stability of households, measured in terms of mortgage payment timeliness, foreclosure, and revolving credit utilization and risk of delinquency. Leveraging quasi random variation in the degree of generosity of servicers tasked with delivering forbearance, I find that the program significantly reduced mortgage delinquency and foreclosure rates, with positive spillovers to revolving credit stability during and after the pandemic. Furthermore, I use an event study approach to analyze heterogeneous policy effects driven by both forbearance exit plans and the duration of relief.

References

Adelino, M., M. A. Ferreira, and M. Oliveira (2024). The heterogeneous effects of household debt relief. *Available at SSRN*.

Aiello, D. J. (2022). Financially constrained mortgage servicers. *Journal of Financial Economics 144*(2), 590–610.

Albuquerque, B. and A. Varadi (2022, February). Consumption Effects of Mortgage Payment Holidays: Evidence during the COVID-19 Pandemic. *IMF Working Papers 2022*(044).

Altunok, F., Y. Arslan, and S. Ongena (2023). Monetary policy transmission with adjustable and fixed rate mortgages: The role of credit supply.

An, X., L. Cordell, L. Geng, and K. Lee (2022). Inequality in the time of covid-19: Evidence from mortgage delinquency and forbearance. *Available at SSRN 3789349*.

Auclert, A., W. Dobbie, and P. Goldsmith-Pinkham (2019, March). Macroeconomic Effects of Debt Relief: Consumer Bankruptcy Protections in the Great Recession. (w25685), w25685.

Calabria, M. (2023). Pandemic mortgage forbearance design: A practitioner's perspective. *Regulation 46*, 32.

Calfas, J. (2021). Close to 40% of us households say they face financial difficulties as covid-19 pandemic continues. *The Wall Street Journal*.

Campbell, J. Y., N. Clara, and J. F. Cocco (2021). Structuring mortgages for macroeconomic stability. *The Journal of Finance 76*(5), 2525–2576.

Cherry, S., E. X. Jiang, G. Matvos, T. Piskorski, and A. Seru (2021, January). Government and Private Household Debt Relief during COVID-19.

Consumer Financial Protection Bureau (2021). COVID-19 Prioritized Assessments, Special Edition.

Consumer Financial Protection Bureau (n.d.). It takes a plan to exit mortgage forbearance. Accessed: [2024-11-01].

Di Maggio, M., A. Kermani, and C. J. Palmer (2020, May). How Quantitative Easing Works: Evidence on the Refinancing Channel. *The Review of Economic Studies 87*(3), 1498–1528.

Dinerstein, M., C. Yannelis, and C.-T. Chen (2023). Debt moratoria: Evidence from student loan forbearance. Technical report, National Bureau of Economic Research.

Dobbie, W., P. Goldsmith-Pinkham, and C. S. Yang (2017). Consumer bankruptcy and financial health. *Review of Economics and Statistics 99*(5), 853–869.

Dobbie, W. and J. Song (2015). Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. *American economic review 105*(3), 1272–1311.

Dobbie, W. and J. Song (2020). Targeted debt relief and the origins of financial distress: Experimental evidence from distressed credit card borrowers. *American Economic Review 110*(4), 984–1018.

Duncan, D. G. (2020, August). Covid-19: The need for consumer outreach and home purchase/financing digitization. *Perspectives Blog*. Executive Advisor, Senior Vice President.

Farrell, D., F. Greig, and C. Zhao (2020). Did mortgage forbearance reach the right homeowners? income and liquid assets trends for homeowners during the covid-19 pandemic. *Income and Liquid Assets Trends for Homeowners during the COVID-19 Pandemic (December 3, 2020)*.

Ganong, P., F. Greig, P. Noel, D. M. Sullivan, and J. Vavra (2022). Lessons learned from expanded unemployment insurance during covid-19. *Recession Remedies: Lessons Learned from the US Economic Policy Response to COVID 19*.

Ganong, P. and P. Noel (2022, October). Why do Borrowers Default on Mortgages?*. *The Quarterly Journal of Economics*, qjac040.

Gerardi, K., L. Lambie-Hanson, P. Willen, et al. (2022). Lessons learned from mortgage borrower policies and outcomes during the covid-19 pandemic. *Federal Reserve Bank of Boston Current Policy Perspectives*.

Guren, A. M., A. Krishnamurthy, and T. J. McQuade (2021). Mortgage design in an equilibrium model of the housing market. *The Journal of Finance 76*(1), 113–168.

Keys, B. J., N. Mahoney, and H. Yang (2023). What determines consumer financial distress? place-and person-based factors. *The Review of Financial Studies 36*(1), 42–69.

Kim, Y. S., D. Lee, T. C. Scharlemann, and J. I. Vickery (2022). Intermediation frictions in debt relief: evidence from cares act forbearance. *FRB of New York Staff Report* (1035).

Lee, S. C. and O. Maghzian (2023). Household liquidity and macroeconomic stabilization: Evidence from mortgage forbearance.

Mae, F. Fhfa announces payment deferral as new repayment option for homeowners in covid-19 forbearance plans.

Mian, A. and A. Sufi (2010, May). The Great Recession: Lessons from Microeconomic Data. *American Economic Review 100*(2), 51–56.

Mian, A. and A. Sufi (2014, May). *House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening Again*. University of Chicago Press.

Patane, C. (2021, January). Servicers report biggest challenges implementing covid-19 assistance programs. *Perspectives Blog*. Vice President, Single-Family Counterparty Risk Oversight.

Understanding Household Financial Decision-Making: Observed Behavior versus Self-Reported Expectations

While the first chapter draws on rich credit bureau data to capture observed financial behavior—offering a clear, grabular view of how households respond to debt relief—our understanding of household expectations and their macroeconomic implications remains more opaque. Despite the widespread use of household inflation expectations in theoretical macroeconomic models, the empirical foundations of these expectations are arguably shakier. Survey data, while abundant, may not reliably reflect how households actually form or act on their beliefs. This tension motivates the second chapter, which investigates how the framing of questions in household surveys can significantly influence reported expectations—particularly for inflation—raising critical concerns about how these widely used measures are interpreted and incorporated into policy analysis.

Chapter 2: On Eliciting Subjective Probability Distributions of Expectations

Valerie Boctor* Olivier Coibion† Yuriy Gorodnichenko‡ Michael Weber§

April 15, 2024

Abstract

Abstract: Using data from a large survey of American households, we compare density forecasts elicited with bins- and scenarios-based questions. We show that inflation density forecasts are sensitive to the survey question designs used to elicit them. The within-person discrepancy is smaller, but still discernible, for unemployment expectations. The discrepancy in responses is systematically related to sociodemographic characteristics of respondents. The differences shed light on the significance of priming in bins-based inflation density forecasts.

JEL Codes: C83, D84, E31
Keywords: Expectations, surveys, inflation, unemployment

*University of California, Berkeley
†University of Texas, Austin and NBER
‡University of California, Berkeley and NBER
§Chicago Booth and NBER

"*How should we measure inflation expectations, and how should we use that information for forecasting and controlling inflation? I certainly do not have complete answers to those questions, but I believe that they are of great practical importance*." Ben Bernanke (2007).

I. Introduction

Expectations are a core element of modern macroeconomic models and policymaking. As a result, measuring expectations is central for a broad spectrum of applications ranging from understanding the Phillips curve to quantifying uncertainty to managing expectations for macroeconomic stabilization. However, eliciting subjective expectations— especially subjective probability distributions— from surveys of households and firms (i.e., the general public) is fraught with a number of measurement issues, such as limited numeric and economic literacy of respondents. Our objective is to compare subjective expectations elicited via different methods to highlight potential differences in outcomes and help researchers and practitioners choose methods appropriate for their applications. Specifically, we focus on two popular survey designs: *i*) the bins design popularized by the Survey of Consumer Expectations which is run by the Federal Reserve Bank of New York (Potter et al. 2017), and *ii*) the scenarios design proposed by Bloom et al. (2020).[1]

Using a large survey of U.S. households presented with questions based on both survey designs, we find that there are significant differences in measured subjective expectations across the designs. For example, scenario-based inflation expectations tend to convey higher levels and greater uncertainty than bins-based inflation expectations. At the same time, the cross-design differences are smaller for unemployment expectations. We also find that bins-based design may result in lumping responses at extreme bins and may prime respondents to choose unlikely outcomes. For example, few households envision deflation when expectations are elicited via scenario-based questions or via point predictions.[2] In contrast, many households assign positive probability to deflation in bin-based questions that include deflation as a possible outcome. We observe that when households are free to choose possible outcomes (especially for inflation), they tend to report scenarios outside the ranges offered in bins-based questions. This pattern reduces consistency across methods for inflation expectations but the discrepancies are smaller for unemployment expectations.

Our paper contributes to several strands of research. First, Manski (2004, 2017) and others discuss the pros and cons of using different methods to elicit subjective probability distributions and measure uncertainty.[3] We provide a novel within-respondent comparison of two leading methods and thus shed new light on their strengths and weaknesses. Within this literature, key recent papers document the sensitivity of survey responses to question designs: Becker et al. (2023) show in online surveys that the location of the midpoint of the proposed probability distribution and the size of the bins shape the implied mean and uncertainty of reported expectations. Hayo and Méon (2023) find that individuals are significantly less likely to form inflation expectations when responding to a free-response question instead of a guided format. Our work advances the survey design sensitivity literature by analyzing the discrepancies that emerge in responses to two

[1] Because of space constraints in our survey, we did not study min-max-midpoint approach popularized by Guiso, Jappelli and Pistaferri (2002).
[2] Gorodnichenko and Sergeyev (2021) document this pattern holds for many advanced economies, including Japan.
[3] See Bruine de Bruin et al. 2023 for a survey of this literature.

guided formats. Furthermore, a large body of work studies demographic predictors of inflation expectations.[4] Our contribution is to document predictors of discrepancies in responses across survey designs.

II. Background and Survey Design

We utilize three methods to elicit subjective expectations. The first method is based on the New York Fed's Survey of Consumer Expectations (SCE). In this influential survey design, respondents are asked to report their subjective probabilities for 10 bins of possible inflation values. The wording of the question is

We would like to ask you some questions about the overall economy and in particular about the rate of inflation/deflation (Note: inflation is the percentage rise in overall prices in the economy, most commonly measured by the Consumer Price Index and deflation corresponds to when prices are falling).

In this question, you will be asked about the probability (PERCENT CHANCE) of something happening. The percent chance must be a number between 0 and 100 and the sum of your answers must add up to 100.

What do you think is the percent chance that, **over the next 12 months**…

	Percentage Chance
the rate of inflation will be 12% or more	______
the rate of inflation will be between 8% and 12%	______
the rate of inflation will be between 4% and 8%	______
the rate of inflation will be between 2% and 4%	______
the rate of inflation will be between 0% and 2%	______
the rate of deflation (opposite of inflation) will be between 0% and 2%	______
the rate of deflation (opposite of inflation) will be between 2% and 4%	______
the rate of deflation (opposite of inflation) will be between 4% and 8%	______
the rate of deflation (opposite of inflation) will be between 8% and 12%	______
the rate of deflation (opposite of inflation) will be 12% or more	______
% Total	______

The main advantage of the bins-based design is that it yields probability distributions for each respondent, as opposed to point forecasts, allowing researchers to infer person-level uncertainty from the probability-weighted dispersion of point values around the implied mean. On the downside, Coibion et al. (2020) document several potential problems related to priming that may arise from offering respondents a fixed grid of possible outcomes. Here and in other designs respondents are asked to forecast a specific price index.

The second method is the scenarios-based design proposed by Bloom et al. (2020). This design asks respondents to report values for low-, medium-, and high-inflation scenarios (in some cases, respondents are asked to provide five scenarios) and assign subjective probabilities to each scenario. The wording of the question is:

[4] See D'Acunto et al. 2023, Weber et al. 2022, and D'Acunto and Weber (forthcoming) for surveys.

Over the next 12 months, which approximate inflation rate (as measured by the Consumer Price Index) would you assign to each of the following scenarios? If you think there was inflation, please enter a positive number. If you think there will be deflation, please enter a negative number. If you think there will be neither inflation nor deflation, please enter zero.

A LOW inflation rate would be about: ______
A MEDIUM inflation rate would be about: ______
A HIGH inflation rate would be about: ______

Please distribute 100 points to the percentage changes you just entered, to indicate how likely you think it is that each inflation rate will happen. The sum of the points you allocate should total to 100.

LOW: The likelihood of realizing a "LOW" inflation rate would be ______
MEDIUM: The likelihood of realizing a "MEDIUM" inflation rate would be ______
HIGH: The likelihood of realizing a ``HIGH'' inflation rate would be
% Total [TOTAL ANSWERS FROM ABOVE – MUST SUM TO 100%] ______

Unlike the bins-based method, this method gives respondents more freedom to pick possible outcomes and thus priming or bunching of responses in extreme bins is less likely. On the other hand, because responses are not supervised, one may obtain a sample with outliers and bunching at multiples of 5. Appendix Figure A1 shows how we convert the three-point responses into probability distributions.

Because bins- and scenario-based questions are cognitively demanding, we also ask respondents to provide their point predictions. The wording of the question is

What do you think the inflation rate (as measured by the Consumer Price Index) is going to be over the <u>next 12 months</u>? Please provide an answer as a percentage change from current prices.

% [RANGE: -100-100, ONE DECIMAL] _____

If you think there will be inflation, please enter a positive number. If you think there will be deflation, please enter a negative number. If you think there will be neither inflation nor deflation, please enter zero.

Similar to the scenario-based method, this question is less likely to prime responses by offering a fixed grid of possible outcomes but is more likely to generate outlier responses. However, we note two important features of this question. First, this wording of the question prompts respondents to contemplate deflation. Second, although this question mimics the Michigan Survey of Consumers (MSC), we do not probe respondents who report high rates of inflation because we want to minimize priming.[5]

[5] The Michigan Survey of Consumers provides this instruction to interviewers, "IF R GIVES AN ANSWER THAT IS GREATER THAN 5%, PLEASE PROBE WITH: 'Let me make sure I have that correct. You said that you expect prices to go (up/down) during the next 12 months by (X) percent. Is that correct?'"

We use the same three methods to elicit expectations for unemployment rate. The wording for the bins-based question is

In THIS question, you will be asked about the probability (PERCENT CHANCE) of something happening. The percent chance must be a number between 0 and 100 and the sum of your answers must add up to 100.

What do you think is the percent chance that, **in 12 months** …

Chance	**Percentage**
Unemployment rate will be more than 20%	______
Unemployment rate will be between 15% and 20%	______
Unemployment rate will be between 10% and 15%	______
Unemployment rate will be between 8% and 10%	______
Unemployment rate will be between 6% and 8%	______
Unemployment rate will be between 4% and 6%	______
Unemployment rate will be less than 4%	______
% Total [TOTAL ANSWERS FROM ABOVE – MUST SUM TO 100%]	______

III. Data

In our empirical analysis, we use the November 2020 wave of the survey that was launched in Coibion, Gorodnichenko and Weber (2022). This survey builds on the Nielsen Homescan Panel which is a popular platform for marketing research. The panel tracks more than 80,000 households who are broadly representative of the U.S. population (sampling weights are provided by Nielsen to correct for any imbalances). The panelists are invited to participate in occasional surveys which typically have a response rate of 20-25%. Participation is rewarded with points which panelists can cash in with Nielsen. Some information about households (e.g., household size and income) is available via Nielsen's background annual surveys and additional information (e.g., current employment status, political leanings) is collected by our survey.

The November 2020 wave targeted electoral issues and the surrounding uncertainty. Given the focus of our analysis, we utilize only questions pertinent to subjective probability distributions of macroeconomic variables: inflation and unemployment. Furthermore, because that survey wave had a randomized controlled trial about elections, we constrain the sample to respondents who were not provided with information in the information treatments (i.e., the control group). We also apply a series of filters to remove noise from survey responses. Specifically, we drop responses that include extreme responses (the point prediction for inflation is greater than 30% or less than -1%, the point prediction for unemployment is greater than 30%) and responses for which a parametric distribution fitting is infeasible, based on the method of Engelberg, Manski and Williams (2009).[6]

[6] We also drop individuals who, in the bins question, report only two, non-adjacent bins, since parametric distribution fitting is not straightforward in those cases. See Appendix B for a detailed description of the parametric approach and robustness.

To provide points of reference, the actual rates for inflation and unemployment at the time of the survey were 1.2% and 6.7%. The Survey of Professional Forecasters, which is run by the Federal Reserve Bank of Philadelphia, projected inflation and unemployment to reach 2.2% and 5.8%. In November 2021, inflation and unemployment rates realized at 6.9% and 4.1%.

IV. Basic moments

In a first pass at the data, we report basic descriptive statistics (raw as well as Huber-robust to outliers) in Table 1. Generally, the distributions of inflation and unemployment expectations have a thick right tail (columns (5)-(9) provide selected percentiles) so that average forecasts (column (1)) exceed medians. In the same spirit, Huber-robust moments (columns (3) and (4)) tend to be moderated relative to moments of the raw data. For example, the average point forecast for inflation is 6.3% in the raw data and 3.6% according to the Huber-robust method. As suggested by Reis (2021), the heavy right tail in inflation expectations in November 2020 turned out to be consistent with high realized inflation in November 2021.

The implied means of inflation expectations from bins-based questions tend to produce lower values than point predictions which in turn tend to be lower than moments implied from scenario-based questions. For instance, Huber-robust averages are 2.5%, 3.6% and 4.1%, respectively. At the same time the ranking is reversed for unemployment expectations.

Irrespective of whether we use point predictions or implied means, there is much disagreement in households' predictions for inflation and unemployment. We find that the cross-sectional standard deviation was about 3.5% for inflation and about 5% for unemployment. At the same time, the average uncertainty (measured by the standard deviation of the provided subjective probability distributions) tends to be lower than disagreement. Although prior work established this pattern for households' and firms' inflation expectations (e.g., Coibion et al. 2021), we are, to the best of our knowledge, the first to document it for unemployment expectations. We also find that the implied uncertainty is smaller in scenario-based distributions than in bins-based distributions. For example, the standard deviations implied by subjective probability distributions for inflation are 4.0% for bins and 1.4% for scenarios (Column 3 of Table 1). The difference is smaller for uncertainty in unemployment expectations, but it is still sizable.

To validate our data, we compare our moments to their counterparts in the SCE and MSC. Broadly, the results are similar across surveys but there are some differences. For instance, we tend to find more disagreement and uncertainty in the Nielsen sample than in the SCE and MSC. Some of this difference should be due to variation in survey designs (e.g., recall that MSC probes responses if reported inflation expectations exceed 5% which could compress the distribution. In our question eliciting point predictions, we ask respondents to report inflation within a value in the range of (-100, 100) *if* they report a value outside this range).

Table 2 provides additional details on the responses to questions eliciting probability distributions for inflation and unemployment. Column (1) shows that roughly 20% of respondents assign positive probability to only one bin and approximately 30% assign positively probabilities to *all* bins. The latter can be surprising given that some bins offer rather extreme scenarios such as deflation of more than 12% (this happened during the recession of 1921) or an unemployment rate greater than 20% (this happened only during the Great Depression). These patterns are consistent with priming of responses (e.g., respondents may feel the need to assign a positive probability to a scenario just because it is offered) and the cognitively demanding nature of the question (e.g., many households can have low financial/numeric literacy and cognition and thus struggle with probabilities, see Lusardi and Mitchell, 2023; D'Acunto et al, 2023).

The scenario-based question also has limitations. Ideally, respondents should have reported three distinct scenarios, each with probability 0< $\Pr\left[\pi^{e}_{i,scen}\right] < 100$. However, we observe repeat scenarios relatively frequently in the data, as well as cases in which one scenario's probability is selected as 100, or the sum of two scenarios' probabilities is 100. As a consequence, we observe that 26% of respondents effectively reported a single value for expected inflation with $\Pr\left[\pi^{e}_{i,scen}\right] = 100$ (column 2). In a similar manner, 16% of individuals effectively report two scenarios for inflation so that only ~60% of respondents effectively provided three scenarios. This issue is less pronounced for unemployment expectations for which more than 80% of respondents provide three scenarios with positive probabilities. Again, these results suggest that low financial and numeric literacy may be a binding constraint and that inflation is a particularly confusing subject for households.[7] However, the potentially unusual choices for bins and scenario are somewhat consistent: those who assign positive probability for a single bin are also more likely to assign positive probability to fewer scenarios.

V. Consistency of subjective expectations

Although different methods to elicit subjective probability distributions yield broadly similar averages and standard deviations, there is dramatic variation across methods in individual responses. Figure 1 presents typical cases for inflation and unemployment expectations. We quantify these differences using several metrics.

A. Endpoint Matching

The first measure is a check of whether the implied distribution supports match exactly. Here, consistency is measured as an indicator variable equal to 1 if the responses from each question imply the same support of possible inflation values, after adjusting scenario values to the nearest corresponding bin value. In particular, for respondent *i*, let LB^{i}_{v} and UB^{i}_{v} denote the extreme values from version $v = \{bins, scen\}$, respectively. Each bin (or scenario) b_n has lower bound b^{L}_{n} and upper bound b^{R}_{n}. Then the support for the reported distribution is given by

$$LB^{i}_{bins} = \min\{ b^{L}_{n} \mid \Pr[\pi^{e}_{i} \in b_n] > 0\}$$

$$UB^{i}_{bins} = \max\{ b^{R}_{n} \mid \Pr[\pi^{e}_{i} \in b_n] > 0\}$$

In principle, consistency implies $LB^{i}_{bins} = LB^{i}_{scen}$ and $UB^{i}_{bins} = UB^{i}_{scen}$, but this comparison is not immediately feasible since the support of the bins- and scenario- based extrema are different. Indeed, the bins effectively run -14% to +14%, whereas scenario range from -100% to 100%. To remedy this, we adjust $(LB^{i}_{scen}, UB^{i}_{scen})$ to the nearest encompassing endpoints from

[7] Bernanke et al. (1999) observed, "Some economists have argued that the public's consistent apathy towards inflation (as evidenced by opinion polls, for example) is primarily the result of confusion about what inflation really is… Somewhat paradoxically, to a degree inflation has become perceived as a serious economic problem precisely because of the public's confusion over what inflation is and about how to make adjustments for it."

the set of possible bin-based extrema, B'. The adjusted values are denoted $\widetilde{LB}^{i}_{scen}$ and $\widetilde{UB}^{i}_{scen}$. In general, each scenario-based value corresponds to exactly one bin, although an exception occurs for certain responses that include only one distinct scenario.[8]

Columns (4)-(6) in Table 2 report results for this consistency check. For inflation expectations, we observe that there is low consistency in the lower bound (approximately 14%). This result obtains because respondents tend to assign positive probability to deflation in bins question but almost never envision deflation in scenario-based questions. The consistency rate is higher for the upper bound (approximately 36%) mainly because many respondents assign positively probability to the top inflation bin and thus nest high inflation scenarios. The consistency for both margins is rare (less than 5%). For unemployment expectations, the results are more similar for upper and lower bounds but the rate of consistency is very low too.

B. Distribution Support Overlap

To provide a sense of the intensive margin for consistency, we consider how often the implied distribution supports from the bins and scenarios questions overlap. For each respondent, we calculate the percent overlap by summing up the number of values that are reported as possible in both versions, then dividing by the total sum of the support ranges in each question version.[9] Specifically, we define the share of overlapping values as

$$O_i = \frac{\min\left\{UB^{i}_{bins}, \widetilde{UB}^{i}_{scen}\right\} - \max\left\{LB^{i}_{bins}, \widetilde{LB}^{i}_{scen}\right\}}{\frac{1}{2}\left[\widetilde{UB}^{i}_{scen} - \widetilde{LB}^{i}_{scen}\right] + \frac{1}{2}\left[UB^{i}_{bins} - LB^{i}_{bins}\right]}$$

where the numerator is the minimum upper bound for scenario- and bin-based responses minus the upper lower bound and the denominator is the average range for both types of questions.

We find (column 7 in Table 2) that the average percent overlap between the values reported in the bins and data is 35% for inflation expectations and 59% for unemployment expectations. We leave it to the reader to decide whether this is a half-full or half-empty glass but the method of eliciting subjective expectations is potentially important.

C. Point forecasts vs. Implied means and uncertainty

Although we do not observe "the true subjective expectations" and thus we cannot have a clear benchmark for validating responses in the distributional questions, one can use point predictions as a benchmark because point predictions are less cognitively demanding and the question design generally has less priming. Panel A of Figure 2 presents binscatters for implied means vs point predictions of inflation expectations. There is a fairly weak relationship between point predictions and bins-based implied means: regressing implied means on point prediction yields an estimated slope of 0.11 (standard error 0.01) and a $R^2 = 0.05$. Clearly, implied means level off for high point predictions. This pattern is consistent with a cap on the maximum expected inflation that

[8] For special cases where single-value responses lie on an endpoint from the overlapping b_ns, the transformation is 1-to-2. E.g., example, suppose $\pi^e_i = 4$. Then there are two corresponding bins that would be deemed "consistent"": either $(LB^i_1, UB^i_1) = (2,4)$ or $(LB^i_1, UB^i_1) = (4,8)$. Both of these are permitted in the consistency check.

[9] This metric assumes that intervals are continuous for both versions. For the bins-based data, this is oversimplifying in cases where respondents report positive probability for bins that are not consecutive. Thus, the overlap rates calculated with our formula can be considered as an upper bound.

respondents can convey (recall that the top bin is inflation of 12% and above which we code as 16%). This pattern is also consistent with priming of responses in that respondents are nudged to consider more moderate outcomes for inflation based on the bins that they see. Because scenario-based questions do not have fixed bins and are less likely to suffer from priming, one should expect a stronger relationship between scenario-based implied mean and point prediction. This prediction is borne out by the data: The slope is now 0.61 (standard error 0.01), still less than 1 but substantially larger than for the subjective distribution with a $R^2$of 0.61. Furthermore, the relationship between implied means for bins-based distributions is stronger for the unemployment rate (Panel B) which is consistent with wider and a less binding set of bins for unemployment. In other words, one may expect more consistency of bins-based implied means if bins cover a wider range of possible outcomes rather than limit them to be between -16% to +16% inflation.

In a related exercise, we examine how uncertainty is related to point predictions. Intuitively, higher inflation is associated with more volatile inflation and thus one may expect a positive relationship between point predictions and uncertainty. Panels C shows that bins-based uncertainty is systematically above scenario-based uncertainty for inflation expectations. Furthermore, there is U-shaped relationship between bins-based uncertainty and point predictions when point predictions are close zero. We conjecture that the spike in uncertainty for low point predictions comes from respondents confusing the concepts of deflation and inflation in the bins-based question. We observed a similar pattern for unemployment expectations (Panel D).

D. Support and uncertainty

Because individual probability distributions are noisy, it is instructive to examine average (across respondents) CDFs for expectations (Panels G and H of Figure 2). For the bins-based CDF, the cumulative probability is set to 0 at $\pi^e = -14$, and it is set to 100 at $\pi^e = 14$.[10] For the bins-based CDFs, this censoring is required at some chosen values, since we do not observe extreme values within the highest and lowest bins. For the scenarios-based CDFs, the cumulative probability is set to 100 at the maximum support value, which we choose to be 50 in order to smooth the right side of the distribution via linear interpolation.

The CDFs corresponding to the bins- and the scenarios-based inflation expectations exhibit a familiar "S" shape with an inflection point around zero but there are important differences.[11] First, the distribution implied by the scenarios-based question lies well below the bins CDF, with roughly 40% of the probability mass in the scenarios CDF corresponding to inflation values above the cutoff midpoint value for the bins CDF, $\pi^e = 14$. In other words, the reported inflation expectations in the scenarios CDF are substantially higher than those implied by the traditional bins-based data both at the aggregate and individual levels. A similar finding applies to unemployment expectations. Second, the SCE's maximal bin, which spans [12, UB], potentially obscures a large portion of what could be the "true" aggregate distribution, given the large average probability of inflation above 12% in the scenarios-based CDF. Third, the left tail of the bins-based distribution lies well above the scenarios CDF. Specifically, the bins CDF suggests that households believe deflation will occur with a probability of up to 33.5%, in contrast with a probability of 1.3% for the same inflation range in the scenarios-based CDF. Perhaps not surprisingly, implied

[10] The conventionally assumed extrema for the bins-based distribution support are {-16, 16}. For the bins-based, CDFs, we plot the implied midpoints of the bins, so {-14, 14} are the effective cut-off values.

[11] The CDFs for our bins-based question and the one implied by the SCE data (for November 2020) are similar, which lends credence to the idea that our data is comparable to the SCE, and that our results apply more generally.

uncertainty is correlated across bins-based and scenario-based questions but the relationship is not linear (Panels E and F).

E. Predictors of discrepancies

What respondent characteristics predict differences across methods eliciting expectations? To answer this question, we regress the absolute value of differences across different measures on sociodemographic variables $\boldsymbol{X}$ such as gender, age, educational attainment, income, political leanings, and employment status:

$$\left|Expectation_i^{measure\ \#1} - Expectation_i^{measure\ \#2}\right| = \boldsymbol{X}_i\boldsymbol{\beta} + error$$

The choice of these variables is informed by earlier research documenting that these characteristics can predict cross-sectional variation in macroeconomic expectations (see D'Acunto et al. 2023 for a survey). For example, women usually have higher inflation expectations, a fact that we reproduce as well (see Appendix Table A1 for regression estimates). We find (Table 3) that some of these variables can predict discrepancies in responses, too. For example, female respondents tend to have large differences not only for forecasts (columns 1-3 and 4-7, for inflation and unemployment, respectively) but also for uncertainty in their forecasts (columns 4 and 8). Higher incomes and college+ education are associated with smaller discrepancies. Other variables can have some predictive power too, but these associations are less robust. Higher incomes and education are likely associated with stronger cognitive abilities and thus weaker inconsistencies in responses (D'Acunto et al, 2023), yet it is not clear why women would have more dissonance in their responses across different types of survey questions.

VI. Concluding remarks

Measuring macroeconomic expectations of households and firms is a difficult task. Time constraints, limited financial and numeric literacy, lack of economic knowledge, present formidable challenges. At the same time, returns to good measurement are very high for positive and normative economics. To this end, we conduct a systematic comparison of two popular methods (bins- vs scenario-based questions) to elicit subjective probabilistic distributions for inflation and unemployment expectations.

We find that elicited subjective expectations are sensitive to which method is used. There are important differences in the first and second moments as well as the support of the elicited distributions, there is limited correlation in responses across methods. Furthermore, these differences appear to vary systematically across respondent characteristics thus indicating that these differences are more than noise in the data. Although we do not have true subjective expectations to benchmark these two methods, our interpretation of the results suggests that scenario-based elicitation could be a better approach because it is less prone to priming and censoring of responses. Furthermore, because this analysis was done prior to the inflation surge 2021-22, it likely understates how large differences in question formulations may be over time, since the bins questions are generally not altered when inflation rates spike and the associated priming effects become more severe. We hope that our analysis will spur more interest and work in this arena.

References

Becker, Christoph, Peter Duersch, and Thomas Eife. 2023. "Measuring Inflation Expectations: How the Response Scale Shapes Density Forecasts." manuscript.

Bernanke, Ben. 2007. "Inflation Expectations and Inflation Forecasting." Board of Governors of the Federal Reserve System. July 10, 2007. https://www.federalreserve.gov/newsevents/speech/bernanke20070710a.htm.

Bloom, Nicholas, Steven J. Davis, Lucia Foster, Brian Lucking, Scott Ohlmacher, and Itay Saporta-Eksten. 2020. "Business-Level Expectations and Uncertainty." SSRN Scholarly Paper. Rochester, NY. https://papers.ssrn.com/abstract=3808466.Ben Bernanke (2007).

Bruine de Bruin, Wändi, Alycia Chin, Jeff Dominitz, Wilbert van der Klaauw, 2023. "Chapter 1 - Household surveys and probabilistic questions," in R. Bachmann, G. Topa, W. van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, pp. 3-31

Coibion, Olivier Yuriy Gorodnichenko, Saten Kumar, and Jane Ryngaert, 2021. "Do You Know that I Know that You Know…? Higher-Order Beliefs in Survey Data," *Quarterly Journal of Economics* 136(3): 1387-1446.

Coibion, Olivier, Yuriy Gorodnichenko, Saten Kumar, and Mathieu Pedemonte, 2020. "Inflation expectations as a policy tool?," *Journal of International Economics* 124(C): 103297

D'Acunto, Francesco, Ulrike Malmendier, and Michael Weber, 2023. "Chapter 5 - What do the data tell us about inflation expectations?" in R. Bachmann, G. Topa, W. van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, pp. 133-161.

D'Acunto, Francesco and Michael Weber, forthcoming. "Why Survey-Based Subjective Expectations are Meaningful and Important," *Annual Review of Economics*.

D'Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber, 2023. "IQ, Expectations, and Choice," *Review of Economic Studies*, 90(5): 2292-2325.

Engelberg, Joseph, Charles F. Manski, and Jared Williams. 2009. "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters." *Journal of Business and Economic Statistics* 27 (1): 30–41.

Gorodnichenko, Yuriy, and Dmitriy Sergeyev, 2021. "Zero Lower Bound on Inflation Expectations," NBER Working Paper 29496.

Guiso, Luigi, Tullio Jappelli, and Luigi Pistaferri. 2002. "An Empirical Analysis of Earnings and Employment Risk," *Journal of Business and Economic Statistics* 20: 241–253.

Hayo, B. and Méon, P.G., 2023. Measuring household inflation perceptions and expectations: The effect of guided vs non-guided inflation questions. *Journal of Macroeconomics*, *78*, p.103558.

Lusardi, Annamaria, and Olivia S. Mitchell. 2023. "The Importance of Financial Literacy: Opening a New Field." *Journal of Economic Perspectives* 37 (4): 137-54.

Manski, Charles, 2004. "Measuring Expectations," *Econometrica* 72(5): 1329-1376.

Manski, Charles, 2017. Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise, Charles F. Manski. in *NBER Macroeconomics Annual* 2017, volume 32, Eichenbaum and Parker, eds.

Potter, Simon, Marco Del Negro, Giorgio Topa, and Wilbert van der Klaauw. 2017. "The Advantages of Probabilistic Survey Questions." Available at https://papers.ssrn.com/abstract=3098648.

Reis, Ricardo, 2021. "Losing the Inflation Anchors," *Brookings Papers on Economic Activity* 52(2 (Fall)), pages 307-379.

Weber, Michael, Francesco D'Acunto, Yuriy Gorodnichenko, and Olivier Coibion. 2022. "The Subjective Inflation Expectations of Households and Firms: Measurement, Determinants, and Implications." *Journal of Economic Perspectives* 36(3): 157-84.

Table 1. Moments of 12 months ahead inflation and unemployment expectations in November 2020

	Raw		Huber robust		Percentiles				
	Mean	St. Dev	Mean	St. Dev	P10	P25	P50	P75	P90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Inflation expectations.									
Nielsen									
Point Forecast	6.3	7.1	3.6	3.5	0.0	1.3	4.0	10.0	18.0
Implied means									
Bins	2.8	5.0	2.5	3.3	-2.2	0.0	2.2	5.8	10.0
Scenarios	11.6	17.1	4.1	3.8	0.0	2.1	4.8	13.5	34.0
Implied st.dev. (uncertainty)									
Bins	4.2	3.5	4.0	3.3	0.0	1.0	3.6	7.7	9.0
Scenarios	3.5	5.4	1.4	1.5	0.0	0.0	1.5	3.9	10.8
Survey of Consumer Expectations									
Point Forecast	6.6	6.6	3.5	2.1	1.9	2.0	4.0	10.0	20.0
Implied mean, bins	4.3	3.9	3.5	2.5	0.2	1.6	3.2	6.0	10.1
Implied st.dev. (uncertainty), bins	2.9	2.5	2.2	1.5	0.0	1.1	2.2	3.9	7.2
Michigan Survey of Consumers									
Point Forecast	4.7	3.3	3.7	2.1	1.0	2.0	4.0	5.0	10.0
Panel B. Unemployment expectations.									
Nielsen									
Point Forecast	11.3	7.4	9.4	5.0	4.0	6.0	10.0	15.0	24.0
Implied means									
Bins	11.2	4.9	10.5	3.8	5.2	7.6	10.5	14.1	17.7
Scenarios	14.9	15.4	8.8	5.2	3.5	6.0	9.3	17.5	35.0
Implied st.dev. (uncertainty)									
Bins	3.3	2.5	3.3	2.4	0.0	1.0	3.3	5.6	6.7
Scenarios	4.3	5.6	2.2	1.9	0.0	1.1	2.4	5.1	11.4

Notes: The sample in panel A is restricted to respondents with point predictions between -1% and 30%. The sample in panel B is restricted to respondent with point predictions between 0% and 30%. Implied uncertainty variables are based on 1% winsorized variances.

Table 2. Consistency in bin- and scenario-based questions

Number of bins with positive probability	Share of responses, %	Distribution by: Scenarios with positive probability, %			Distribution by: Share with consistent bounds, %			Average Overlap, %
		One	Two	Three	Lower	Upper	Both	
	(1)	(2)	(3)	(4)	(4)	(5)	(6)	(7)
Panel A: Inflation expectations								
1 Bin	20.1	41.7	15.4	42.9	25.8	22.9	9.2	33.8
2	10.0	20.0	20.0	59.9	32.9	28.3	13.9	51.7
3	8.3	15.0	18.5	66.5	23.7	35.9	10.8	58.8
4	8.0	18.7	14.8	66.6	17.2	34.9	3.2	53.4
5	9.7	18.0	14.2	67.8	20.5	35.5	2.4	55.5
6	3.4	19.3	13.6	67.1	0.9	31.7	0.0	45.3
7	3.1	17.6	15.1	67.2	0.1	38.2	0.1	36.8
8	3.1	19.2	15.8	65.0	0.0	44.6	0.0	29.7
9	4.8	25.5	10.7	63.7	0.2	41.2	0.2	20.0
10 Bins	29.3	26.8	15.4	57.8	0.5	47.7	0.1	11.8
All observations	100.0	25.9	15.7	58.4	14.0	36.3	4.6	34.6
Panel B: Unemployment expectations								
1 Bin	17.8	0.7	24.2	75.2	10.9	12.4	7.0	35.3
2	12.4	0.6	21.7	77.7	15.0	22.0	10.0	61.0
3	11.1	0.3	18.5	81.2	19.8	17.4	14.5	67.9
4	10.6	0.9	13.9	85.2	11.0	12.5	5.0	63.9
5	9.1	0.4	13.4	86.2	9.5	8.6	2.0	65.1
6	9.0	0.2	12.5	87.3	7.4	6.2	5.0	63.1
7 Bins	30.0	0.1	12.8	87.2	0.7	5.0	0.9	59.0
All observations	100.0	0.4	16.8	82.8	8.9	11.0	5.5	57.5

Notes: Column (1) show the share of respondents assigning positive probability to a given number of bins. Columns (2)-(4) show the share of responses with a given number of distinct scenarios. For each row, columns (2)-(4) sum up to 100. Columns (4)-(6) show the share of respondents who report consistent bounds as described in Section V.A. Column (7) report the percent overlap in supports of the reported subjective bins-based and scenario-based distributions. See Section V.B for the method to compute the overlap.

Table 3. Predictors of differences in expectations.

	Inflation				Unemployment			
	Level			Std.	Level			Std.
Measure 1	Point	Point	Bins	Bins	Point	Point	Bins	Bins
Measure 2	Bins	Scenarios	Scenarios	Scenario	Bins	Scenarios	Scenarios	Scenarios
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	1.06***	0.70***	1.29***	0.72***	0.51***	0.86***	0.97***	0.50***
	(0.07)	(0.06)	(0.10)	(0.06)	(0.05)	(0.08)	(0.08)	(0.05)
Age	0.07***	0.03**	0.05**	-0.05***	-0.03**	-0.02	-0.00	-0.01
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
$Age^2/100$	-0.05***	-0.03**	-0.05**	0.03**	0.02**	0.02	-0.01	0.01
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Log Income	-0.70***	-0.32***	-1.00***	-1.19***	-0.86***	-0.82***	-1.34***	-0.88***
	(0.10)	(0.08)	(0.14)	(0.08)	(0.07)	(0.11)	(0.12)	(0.06)
Republican	0.08	0.06	0.06	0.08	0.02	0.06	0.10	0.04
	(0.08)	(0.07)	(0.11)	(0.07)	(0.06)	(0.09)	(0.09)	(0.05)
Green Party	0.39	-0.44	0.26	-1.04**	0.03	-0.57	0.79	0.13
	(0.47)	(0.34)	(0.75)	(0.40)	(0.42)	(0.62)	(0.63)	(0.35)
Libertarian Party	0.05	0.10	0.01	-0.16	-0.12	-0.78***	-0.50*	-0.29**
	(0.26)	(0.21)	(0.33)	(0.19)	(0.18)	(0.23)	(0.26)	(0.15)
Other Party	0.08	0.02	0.21*	0.21***	0.12*	0.06	0.09	0.01
	(0.09)	(0.08)	(0.12)	(0.08)	(0.07)	(0.10)	(0.10)	(0.06)
Party not reported	0.36***	0.08	0.83***	0.87***	0.98***	0.64***	1.21***	0.48***
	(0.13)	(0.11)	(0.18)	(0.11)	(0.10)	(0.16)	(0.17)	(0.09)
Some high school	-0.81***	0.16	0.42	0.79***	0.11	-0.09	0.77*	0.87***
	(0.28)	(0.28)	(0.46)	(0.29)	(0.31)	(0.41)	(0.44)	(0.21)
Graduated high school	0.03	-0.18**	0.09	0.30***	0.12	-0.31***	-0.16	0.09
	(0.10)	(0.09)	(0.14)	(0.09)	(0.07)	(0.12)	(0.12)	(0.07)
Some college	0.00	-0.02	0.17	0.09	-0.04	-0.09	0.07	-0.03
	(0.09)	(0.07)	(0.12)	(0.07)	(0.06)	(0.10)	(0.10)	(0.05)
Post college graduate	-0.41***	-0.39***	-0.68***	-0.42***	-0.11	-0.39***	-0.30***	-0.16***
	(0.10)	(0.08)	(0.12)	(0.08)	(0.07)	(0.11)	(0.11)	(0.06)
Under 30 hours of work	-0.20*	0.26**	-0.06	0.06	0.07	-0.10	0.02	0.13*
	(0.12)	(0.11)	(0.16)	(0.10)	(0.09)	(0.13)	(0.14)	(0.08)
30-34 hours of work	0.11	0.27**	0.22	0.41***	-0.03	0.50**	0.01	0.18*
	(0.16)	(0.13)	(0.23)	(0.14)	(0.12)	(0.21)	(0.19)	(0.11)
Not employed for pay	-0.15*	0.12	-0.09	0.16**	-0.17***	-0.15	-0.27***	-0.08
	(0.09)	(0.08)	(0.12)	(0.08)	(0.06)	(0.10)	(0.10)	(0.06)

Observations	7,361	6,651	6,891	7,620	6,867	6,502	6,511	6,875
R-squared	0.04	0.03	0.05	0.13	0.08	0.04	0.08	0.09

Notes: The table reports estimates for regressions of absolute value of difference in expectations between two methods (measure #1 and Measure #2) on sociodemographic characteristics of responses. All specifications are estimated using Huber robust regression. ***,**,* denote statistical significance at 1, 5, and 10 percent levels.

Figure 1. Distribution of subjection probabilistic distributions for selected respondents.

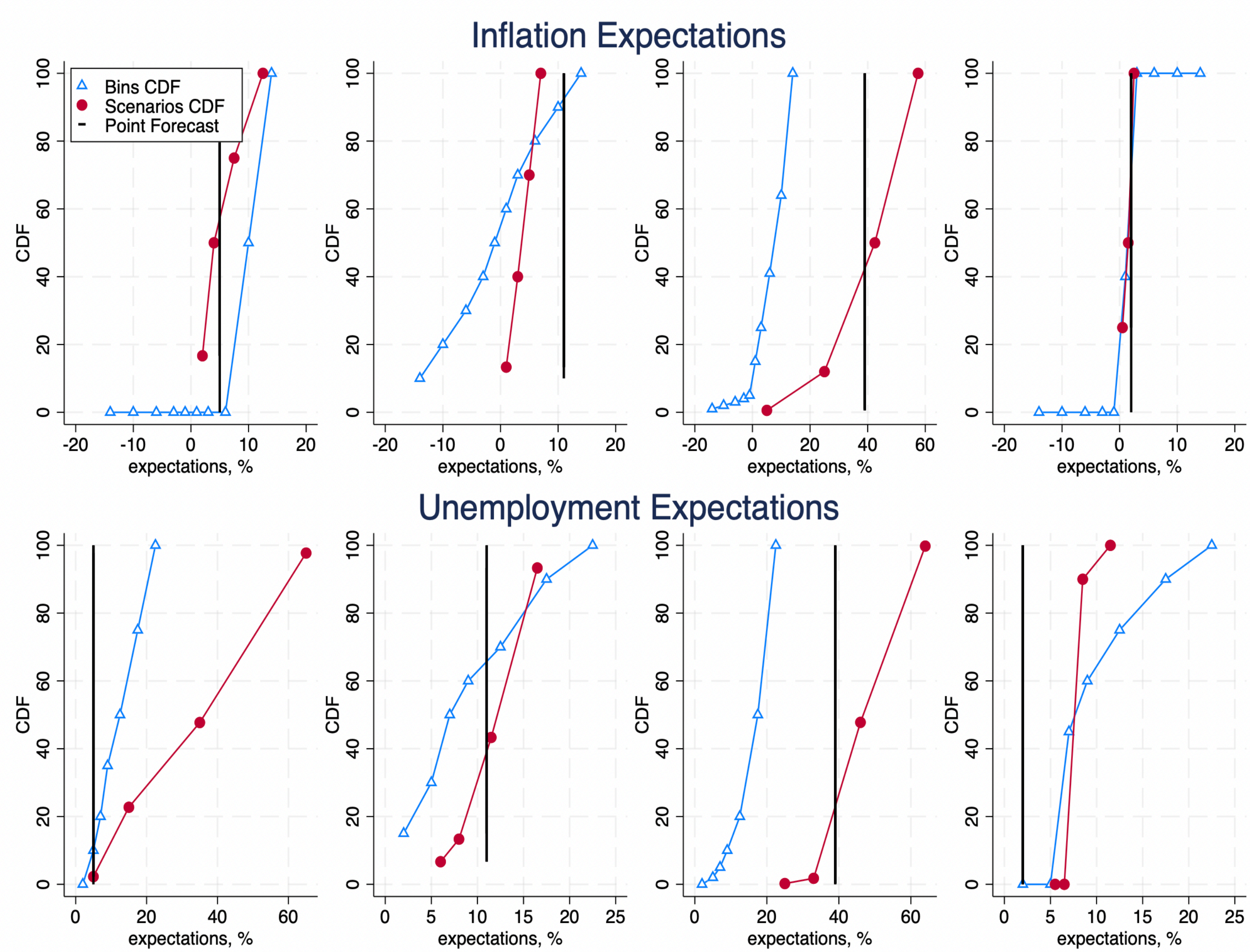
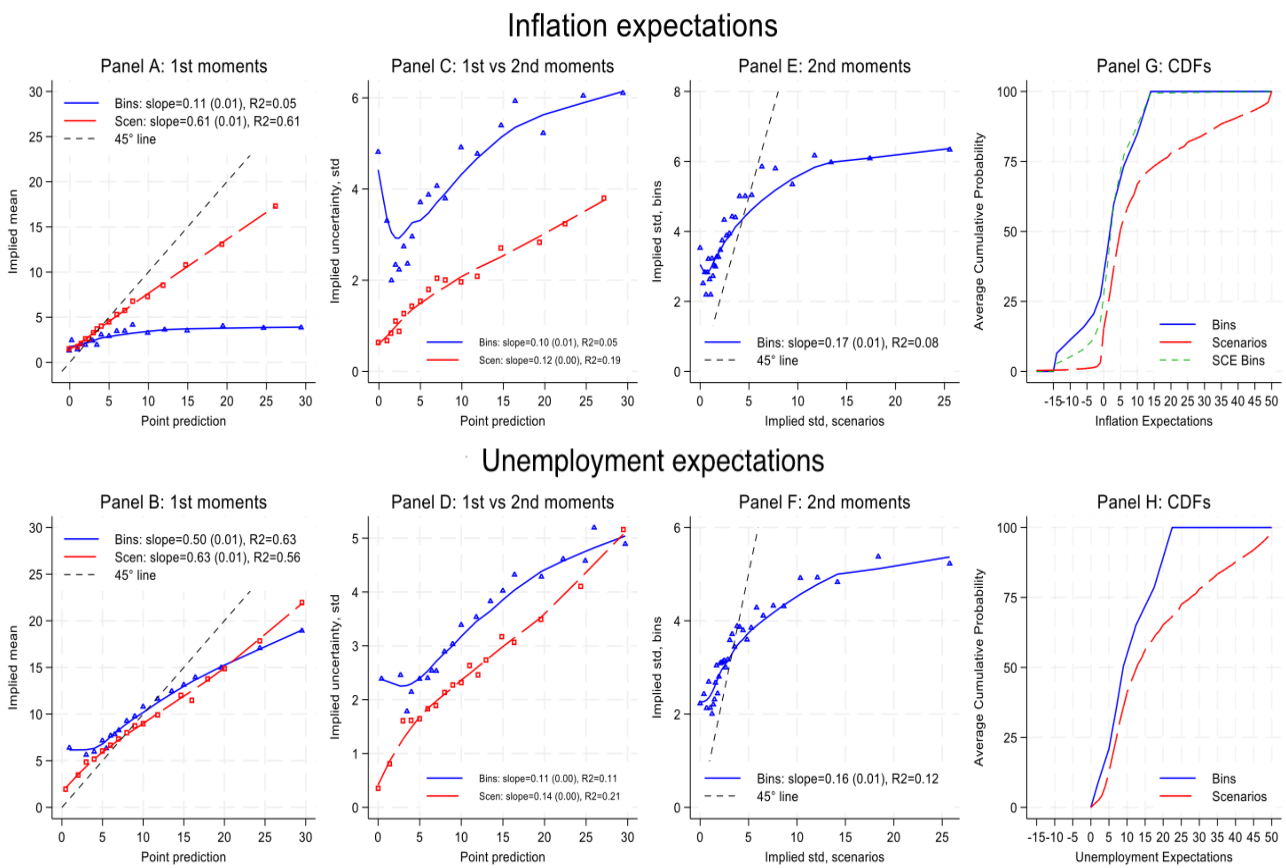


Figure 2. Comparison of moments across methods.

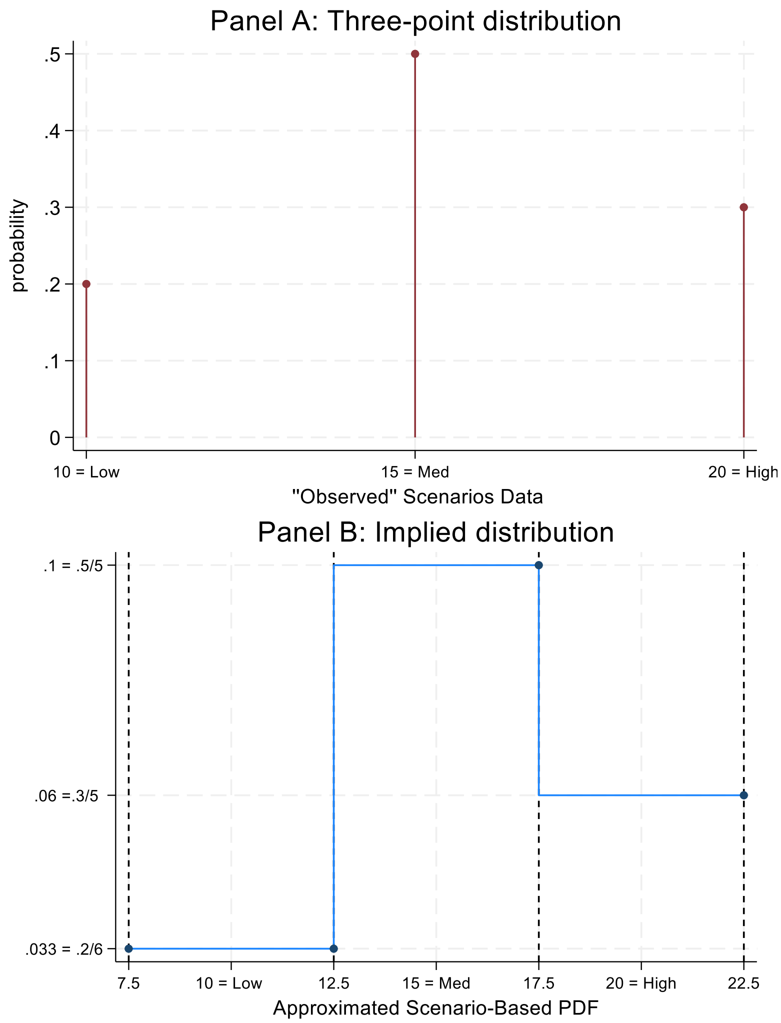


Notes: Panels A-F present binscatter plots. Panels G and H report average (across respondents) cumulative distribution functions (CDFs). Blue solid line is from bins-based elicitation. Red, long-dash line is from scenario-based elicitation. Green, dashed line (Panel G) is from bins-based elicitation in the Survey of Consumer Expectations. Black, short-dash line is the 45° line.

Online Appendix A

Additional Figures and Tables

Appendix Figure A1. Stylized Example of PDF Interpretation for Scenarios Data.



The left panel shows a stylized example of reported expected inflation values and corresponding probabilities for low- medium- and high-inflation scenarios. The right panel demonstrates the method we use to interpret the data as a PDF. In principle, we set each scenario to be the midpoint of a uniformly distributed range of values around that point. The probability within each range is given by the corresponding scenario probability, divided by the number of discrete points in the range of values. As shown above, if an individual reports three scenarios, this results in a 4- point mapping required to pin down the approximated PDF (and CDF). The method is analogous for individuals who report two scenarios, i.e., the approximated distribution has a 3-point mapping. For individuals that report a single scenario with 100% probability, we interpret the CDF as-is.

Appendix Table A1. Predictors of expectations.

	Inflation					Unemployment				
	Level			Uncertainty		Level			Uncertainty	
	Point	Bins	Scenarios	Bins	Scenarios	Point	Bins	Scenarios	Bins	Scenarios
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	0.77***	0.64***	0.93***	0.83***	0.28***	1.87***	1.26***	1.52***	0.55***	0.58***
	(0.09)	(0.08)	(0.10)	(0.08)	(0.04)	(0.12)	(0.10)	(0.13)	(0.07)	(0.05)
Age	0.08***	0.08***	0.07***	-0.09***	0.02***	0.02	-0.04**	0.04	-0.07***	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.01)	(0.01)
$Age^2/100$	-0.04**	-0.05***	-0.03*	0.05***	-0.01	-0.02	0.01	-0.02	0.04***	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)
Log Income	0.10	0.19**	-0.43***	-1.45***	-0.02	-0.18	-0.60***	-0.33*	-0.99***	0.05
	(0.10)	(0.09)	(0.13)	(0.09)	(0.05)	(0.17)	(0.11)	(0.19)	(0.07)	(0.07)
Republican	0.40***	0.38***	0.27**	-0.27***	0.02	-0.35**	-0.31***	-0.09	-0.28***	0.07
	(0.10)	(0.09)	(0.11)	(0.09)	(0.04)	(0.14)	(0.11)	(0.15)	(0.07)	(0.06)
Green Party	0.56	0.85	0.60	0.23	0.57*	-1.23	0.06	-0.75	0.27	0.60
	(0.56)	(0.67)	(0.62)	(0.59)	(0.31)	(1.14)	(0.74)	(1.14)	(0.55)	(0.40)
Libertarian Party	0.71**	1.43***	0.01	-0.71***	0.04	-0.42	-0.49	-0.79*	-0.82***	-0.27*
	(0.32)	(0.31)	(0.29)	(0.25)	(0.14)	(0.44)	(0.35)	(0.41)	(0.22)	(0.16)
Other Party	0.06	0.13	0.08	0.06	-0.03	0.36**	0.23*	0.13	-0.02	0.08
	(0.11)	(0.10)	(0.12)	(0.10)	(0.05)	(0.16)	(0.12)	(0.17)	(0.08)	(0.07)
Party not reported	-0.44***	-0.27**	-0.21	0.90***	-0.21***	-0.42*	0.86***	-0.80***	0.44***	-0.31***
	(0.15)	(0.13)	(0.17)	(0.14)	(0.07)	(0.23)	(0.16)	(0.25)	(0.11)	(0.09)
Some high school	-1.28***	-1.65***	-0.50	2.30***	-0.13	-2.94***	-0.08	-2.09***	1.08***	-0.39
	(0.33)	(0.27)	(0.45)	(0.29)	(0.19)	(0.55)	(0.43)	(0.79)	(0.28)	(0.26)
Graduated high school	-0.32***	-0.67***	-0.72***	0.78***	-0.35***	-0.82***	-0.02	-1.06***	0.40***	-0.50***
	(0.12)	(0.11)	(0.14)	(0.11)	(0.05)	(0.18)	(0.13)	(0.19)	(0.09)	(0.07)
Some college	-0.03	0.11	0.02	0.13	-0.05	0.15	0.33***	-0.16	0.01	-0.18***
	(0.10)	(0.10)	(0.12)	(0.09)	(0.05)	(0.15)	(0.12)	(0.16)	(0.07)	(0.06)
Post college graduate	-0.33***	-0.12	-0.64***	-0.53***	-0.10*	-0.31*	-0.25*	-0.29*	-0.06	-0.17**
	(0.11)	(0.11)	(0.12)	(0.11)	(0.05)	(0.17)	(0.14)	(0.17)	(0.09)	(0.07)
Under 30 hours of work	-0.45***	-0.04	-0.14	0.01	0.04	0.22	-0.02	-0.08	0.06	0.06
	(0.13)	(0.13)	(0.16)	(0.12)	(0.06)	(0.20)	(0.16)	(0.22)	(0.10)	(0.09)
30-34 hours of work	-0.33*	-0.01	-0.10	0.33**	0.12	0.80***	0.10	0.66**	0.24*	0.09
	(0.18)	(0.18)	(0.21)	(0.17)	(0.09)	(0.30)	(0.21)	(0.32)	(0.13)	(0.11)
Not employed for pay	-0.11	-0.11	-0.35***	-0.00	-0.10**	0.04	-0.13	0.32*	0.05	0.00
	(0.11)	(0.09)	(0.12)	(0.10)	(0.05)	(0.16)	(0.12)	(0.17)	(0.08)	(0.06)
Observations	7,386	7,677	6,790	7,845	7,103	7,124	7,237	6,618	7,289	6,701
R-squared	0.04	0.04	0.03	0.17	0.02	0.04	0.06	0.03	0.12	0.03

Notes: The table reports estimates for regressions of expectations on sociodemographic characteristics of responses. All specifications are estimated using Huber robust regression. ***,**,* denote statistical significance at 1, 5, and 10 percent levels.

Appendix B

Construction of CDFs

Naturally, one could wonder how sensitive our results are to the methods we use to interpret data as PDFs. Specifically, in our handling of the scenarios data, we assume a particular method for interpreting individual PDFs (see Appendix A Figure 1), choose the support of the aggregate distribution as [-20, 50], and use linear interpolation to smooth the right side of the aggregate distribution. At the same time, we take a conventional approach when analyzing the bins-based data, using the NY Fed's implicit range of [-16, 16] as the distribution support, and assuming the probability within each bin is distributed uniformly. To ensure that the key features of the bins- and scenarios- based aggregate CDFs are robust to minor differences in data treatment, we abstract from these empirical choices, to the extent possible, and instead use the parametric approach of Engelberg, Manski, and Williams (2009) (hereafter referred to as EMW). After fitting the individual PDFs to continuous parametric distributions, we show that the implied aggregate CDFs are strikingly similar to those in our main results (see Figure B1). Below, we describe the EMW method, including some minor adaptations to fit the scenarios data, and discuss the aggregate CDF results.

In the EMW method, individual bins-based probability data are fitted to parametric distributions using a small set of assumptions on the parameters and estimation with non-linear least squares, as needed. The target parametric distribution for each response is determined by the number of bins (or scenarios) used. One- and two-bin (scenario) responses are fitted to the uniform and isosceles triangular distributions, respectively.[12] Responses including three or more bins (or three scenarios) are fitted to the generalized beta distribution.

For any distribution, we denote the set of parameters as $\theta = \{\eta, l\, s\}$. l is the location parameter, equivalent to the left endpoint of the support, and $s = r - l$ is a scale parameter, equal to the distance between the right and left endpoints, and η is a set of shape parameters, whose elements depend on the specific distribution. For the beta distribution, $\eta = \{\alpha, \beta\}$. For the isosceles triangular distribution, $\eta \equiv .5$ by definition of isosceles. Finally, since the uniform distribution does not take a shape parameter, $\eta = \{\cdot\}$. We describe each of the parametric distributions, using the notation stated above.[13]

<u>Generalized Beta Distribution</u>

$$\eta = \{\alpha, \beta\}; \theta = \{\{\alpha, \beta\}, l, s\}$$

The probability distribution function is given by:

$$f(x;\theta) = \begin{cases} 0, & x < l \\ \dfrac{(x-l)^{\alpha-1}(l+s-x)^{\beta-1}}{B(\alpha,\beta)s^{\alpha+\beta-1}}, & l \le x \le l+s \\ 0, & x > l+s\,, \end{cases}$$

where $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

[12] Responses including only two bins which are not adjacent are not fitted to any parametric distribution. Due to this, these observations are omitted from the sample throughout the analysis in this paper.

[13] Our PDF notation differs from Armantier (2017) and EMW (2009) and reflects our use of the SciPy library in Python to estimate the default "location" and "scale" parameters (as opposed to estimating left and right endpoints).

As stated in EMW, one of the main advantages of using the generalized beta distribution is that it is highly flexible, by dint of the two shape parameters. By the same token, the low number of data points (between 3 to 10) per response relative to the number of parameters (up to four) implies that estimating $\widehat{\theta_i}$ precisely is a challenge, and estimates are notably sensitive to the initial guess, θ_0, as a result. Furthermore, in the bins data, it is relatively rare to observe $\widehat{r_i} \gg r_0$ or $\widehat{l_i} \ll l_0$. We attribute the lack of substantial variation in the endpoint estimates to compression of the underlying bins data, related to priming.

Isosceles Triangular Distribution

$$\eta = \{.5\}; \theta = \{\{.5\}, l, s\}$$

The probability distribution function is given by:

$$f(x;\theta) = \begin{cases} \dfrac{4(x-l)}{s^2}, & l \leq x \leq \dfrac{s+2l}{2} \\ \dfrac{4}{s^2}(s+l-x), & \dfrac{s+2l}{2} < x \leq s+l \\ 0, & otherwise \end{cases}$$

Uniform Distribution

$$\eta = \{\cdot\}; \theta = \left\{\{\cdot\}, l, s\right\}$$

The probability distribution function is given by:

$$f(x;\theta) = \begin{cases} \dfrac{1}{s}, & l \leq x \leq s+l \\ 0, & otherwise \end{cases}$$

Whenever the uniform distribution is used, we assume $l_i = \overline{l_i}$ and $r_i = \bar{r}_i$, which together imply $s_i = \overline{s_i}$. Since all parameters in θ are known, no estimation is required.

After classifying responses by parametric distribution type, we split them into cases depending on which, if either, of the distribution's support endpoints are known, given the assumptions in the EMW method. For any given distribution, the general endpoint cases are defined as follows. (See Table B2 for additional information on how we assign support ranges according to distribution type and endpoint case.)

- Case 0: l, r unknown; $\theta = \{\eta, l, s\}$
  This case describes PDFs for which neither endpoint of the support is known. Both l, s are estimated using non-linear least squares.

  Responses are classified as Case I if:

  Bins: the respondent uses 3+ bins, including the largest and smallest bins.

  Scenarios: 3 distinct scenarios, each with positive probability are used.

- Case 1: l unknown, $r = \bar{r}$; $\theta = \{\eta, l\}$
  This occurs when the max value of the support is pinned down by the data, based on the method of EMW. In this case estimating l is sufficient for s, since $s = \bar{r} - l$.

  Responses are classified as Case II if:

  Bins: the respondent uses at least 3 bins, including the lowest (left-censored) bin, but excluding the highest (right-censored) bin; or, the respondent uses exactly 2 bins, the higher bin used is not censored, and has higher probability than the lower bin.

  Scenarios: the respondent uses 2 distinct scenarios with positive probabilities. In addition, the higher scenario value is assigned a higher probability.

- Case 2: r unknown, $l = \bar{l}$; $\theta = \{\eta, s\}$
  This occurs when the min value of the support is pinned down by the data, based on the method of EMW. In this case estimating s is sufficient to recover r , since $r = \bar{l} + s$.
  Responses are classified as Case III if:

  Bins: the respondent uses at least 3 bins, excluding the lowest (left-censored) bin, but including the highest (right-censored) bin; or, the respondent uses exactly 2 bins, the lower bin used is not censored, and has higher probability than the higher bin.

  Scenarios: the respondent uses 2 distinct scenarios with positive probabilities. In addition, the lower scenario value is assigned a higher probability.

- Case 3: $l = \bar{l}, r = \bar{r}$; $\theta = \{\eta\}$
  This case describes situations in which both endpoints are pinned down by the observed data, based on the assumptions in EMW. We estimate only η, if needed. Recall that $\eta \equiv .5$ for the isosceles triangular distribution, and $\eta = \{.\}$ for the uniform distribution, so estimation is required only if the target distribution is the generalized beta distribution.

  Responses are classified as Case IV if:

  Bins: the respondent uses exactly 1 bin.

  Scenarios: the respondent uses only 1 distinct scenario with 100% probability, or the respondent uses two bins, each having 50% probability.

After fitting each response to the corresponding PDF type, we use the resulting parameter estimates to obtain micro-level CDF values across the grid ranging [-20, 40]. In estimation, the location (left endpoint) parameter is bounded below at -100, and the scale parameter is unconstrained. If applicable, the shape parameters are also unconstrained. We find that the estimation results are relatively invariant to alternative assumptions on the bounds. Using the parametric approach delineated above, we show that the implied aggregate CDFs for both the bins and scenarios subsamples track closely with the original CDFs used in our main results. One noteworthy difference is that the tails of parametric scenarios CDF are flatter than those of the corresponding non-parametric curve, which implies the extreme values of the support could lie beyond [-20,40] This finding bolsters our view that the conventional cut-off values of ±16 for the bins-based distributions are unrealistic. Overall, the similarity between the aggregate parametric and non-parametric curves provide evidence that our results to obtain even when using an alternative set of assumptions over the underlying data. This reinforces the validity of our original approach and demonstrates the robustness of our main results.

Figure B1: Comparison of Parametric and Non-Parametric Inflation CDFs

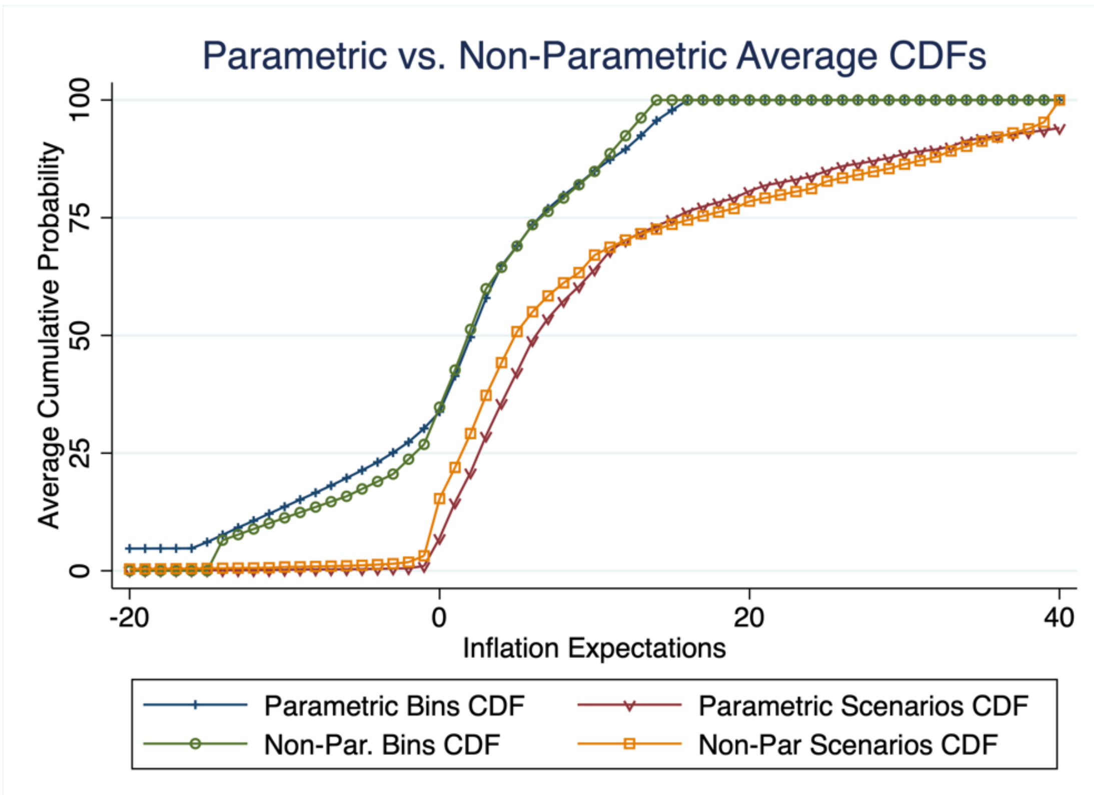


Table B1: Count of Responses by Parametric Distribution Type and Endpoint Case

# Bins/ Scenarios	Case 0	Case 1	Case 2	Case 3
1 (Uniform)				
Bins	0	0	0	1840
Scenarios	0	0	0	2347
2 (Triangular)				
Bins	0	244	237	384
Scenarios	0	409	682	322
3+ (Beta)				
Bins	3129	1595	307	1119
Scenarios	0	0	0	5095
N = 8855				

Table B2: Guide to Assigning Support by Parametric Distribution Type and Endpoint Case

Distr. Type	Case I	Case II	Case III	Case IV
Uniform	N/A	N/A	N/A	Bins: If $\bar{l}_i$ ($\bar{r}_i$) censored, use -16 (16). Scen: Support is $[\pi_i^e - \delta, \pi_i^e + \delta]$, $\delta \rightarrow 0$. No estimation required.
Triangular	N/A	$\bar{r}_i = UB^i_{bins;\, scen}$ $\hat{s}_i \equiv \bar{r}_i - \hat{l}_i$ Estimate l_i.	$\bar{l}_i = LB^i_{bins}$ $\hat{r}_i \equiv \hat{s}_i - \bar{l}_i$ Estimate s_i.	Support is $[LB^i_{bins;\, scen}, UB^i_{bins;scen}]$ No estimation required.
Beta	Endpoints unknown. Estimate $\eta_i\ l_I, s_i$.	$\bar{r}_i = UB^i_{bins;\, scen}$ $\hat{s}_i \equiv \bar{r}_i - \hat{l}_i$ Estimate η_i, l_i.	$\bar{l}_i = LB^i_{bins;\, scen}$ $\hat{r}_i \equiv \hat{s}_i - \bar{l}_i$ Estimate η_i, s_i.	Support is $[LB^i_{bins;scen}, UB^i_{bins;scen}]$ Estimate only η_i.

The second chapter highlights how insights from applied microeconomics—such as attention to survey design and individual-level heterogeneity—can sharpen our understanding of macroeconomic data, particularly expectations. This methodological orientation carries into the third chapter, which applies microeconometric tools to a longstanding macroeconomic question: under what conditions do fiscal deficits generate inflationary risk? By combining a quantile regression technique to estimate an empirical Phillips curve with an inflation-at-risk framework, the final chapter demonstrates how the empirical toolkit of applied microeconomics can be used to rigorously assess macroeconomic regimes and forecast tail risks. In doing so, it offers a cohesive interpretation of COVID-era inflation dynamics in advanced economies—linking observed data, institutional frameworks, and the consequences of large-scale fiscal stimulus during macroeconomic downturns.

Chapter 3: Fiscal deficits and inflation risks: the role of fiscal and monetary regimes*

Ryan Banerjee[†] Valerie Boctor[‡] Aaron Mehrotra[§] Fabrizio Zampolli[¶]

July 1, 2022 (revised May 18, 2025)

Abstract

Using data from a panel of advanced economies over four decades, we show that the inflationary effect of fiscal deficits crucially depends on the prevailing fiscal-monetary policy regime. Under a fiscally-led regime, defined as a regime in which the government does not adjust the primary balance to stabilise debt and the central bank is less independent or puts less emphasis on price stability, the average effect on inflation of higher deficits is found to be up to five times larger than under a monetary-led regime. Under a fiscally-led regime, higher deficits also increase the dispersion of possible future inflationary outcomes, especially the probability of high inflation. Based on forecasts from our model, the high inflation experienced by many countries during the recovery from the Covid-19 pandemic appears more consistent with a fiscally-led regime than a monetary-led regime.

JEL Codes: E31; E52; E62; E63.
Keywords: Fiscal deficit, inflation, fiscal policy regime, monetary policy regime, monetary policy independence.

*We are grateful for comments, suggestions, and discussions to David Archer, Agustín Carstens, Yuriy Gorodnichenko, Luiz Pereira and Hyun Song Shin, as well as to seminar and conference participants at the BIS, University of California at Berkeley, CEF Conference in Dallas, BIS-SNB Workshop in Basel, 50th OeNB Economics Conference and 60th SUERF Anniversary Conference in Vienna, ZEW Public Finance Conference in Mannheim and ABFER 10th Annual Conference in Singapore. We also thank Berenice Martinez for excellent research assistance. The views expressed in the paper are those of the authors and do not necessarily reflect those of the Bank for International Settlements.

†Corresponding author. Bank for International Settlements, Centralbahnplatz 2, 4002 Basel, Switzerland. E-mail address: ryan.banerjee@bis.org

‡University of California, Berkeley. E-mail address: valerieboctor@econ.berkeley.edu

§Bank for International Settlements. E-mail address: aaron.mehrotra@bis.org

¶Bank for International Settlements. E-mail address: fabrizio.zampolli@bis.org

1 Introduction

Governments unleashed a massive wave of fiscal stimulus in the wake of the Covid-19 pandemic, pushing up government debt by around 15 percentage points of global GDP between 2019 and 2021. These developments have triggered a debate about the inflationary consequences of fiscal policies (e.g. Krugman (2021), Summers (2021)). The Economist (2021) has asked whether inflation is a fiscal phenomenon. And some commentators have even argued that the need to manage the high levels of public debt might result in fiscal dominance over monetary policy, posing a risk to price stability (e.g. Landau (2021)).

The theoretical literature has long considered the central role of fiscal policy for inflation. In their seminal paper, Sargent and Wallace (1981) demonstrated the impotence of monetary policy to control inflation when the government runs large fiscal deficits not ultimately financed by taxation. Leeper (1991) argued that the price level would adjust to re-establish the government's intertemporal budget constraint if fiscal policy is unsustainable.

We contribute to this literature by showing that in historical data, the fiscal deficit-inflation relationship crucially depends on the prevailing fiscal-monetary policy regime. To investigate the contribution of fiscal policy to inflation, we estimate an open economy Phillips curve augmented with the fiscal balance using data from 21 advanced economies over four decades.

What distinguishes our analysis from previous research is the careful consideration of the fiscal-monetary policy regime in place. To classify those regimes we use both *de facto* and *de jure* measures. As to the fiscal regimes, our *de facto* classification is based on the result from the seminal paper by Bohn (1998) which shows that fiscal policy satisfies the government's intertemporal budget constraint when the primary surplus is an increasing function of the level of debt relative to GDP. In our baseline specification we follow Mauro et al. (2015) who operationalise Bohn (1998) by estimating fiscal reaction functions in a panel of economies. The authors classify the regimes as either "prudent" or "profligate" depending on whether the estimated response of fiscal surpluses is increasing in the level of debt within a given window. We also consider *de jure* classifications of fiscal regimes based on whether the economy has in place a fiscal rule for the budget balance - that is, legal numerical limits on the overall balance, the structural or cyclically adjusted balance, or the balance over the cycle.

As to monetary policy regimes, we look at whether monetary policy acts to maintain price stability. Our *de jure* classification of monetary policy regimes is based on whether a central bank is classified as being highly or weakly independent. Cukierman (1992) and Cukierman et al. (1992) show that the degree of central bank independence is negatively correlated with inflation in advanced economies. The specific measures we use come from Romelli (2022) who follows the entire set of legislative changes to laws concerning the central bank. Given the particular importance of monetary accommodation of fiscal policy, our baseline *de jure* measure is based on specific limitations on central bank lending to the public sector enshrined in central bank laws. Our *de facto* measure of the monetary policy regime is based on whether the central bank's policy interest rate is below that suggested by a Taylor rule (Taylor (1993)). A number of studies have shown that failure to satisfy the Taylor principle, i.e. not adjusting interest rates by more than with inflation, resulted in inflation instability in the United States during the 1960s and 1970s

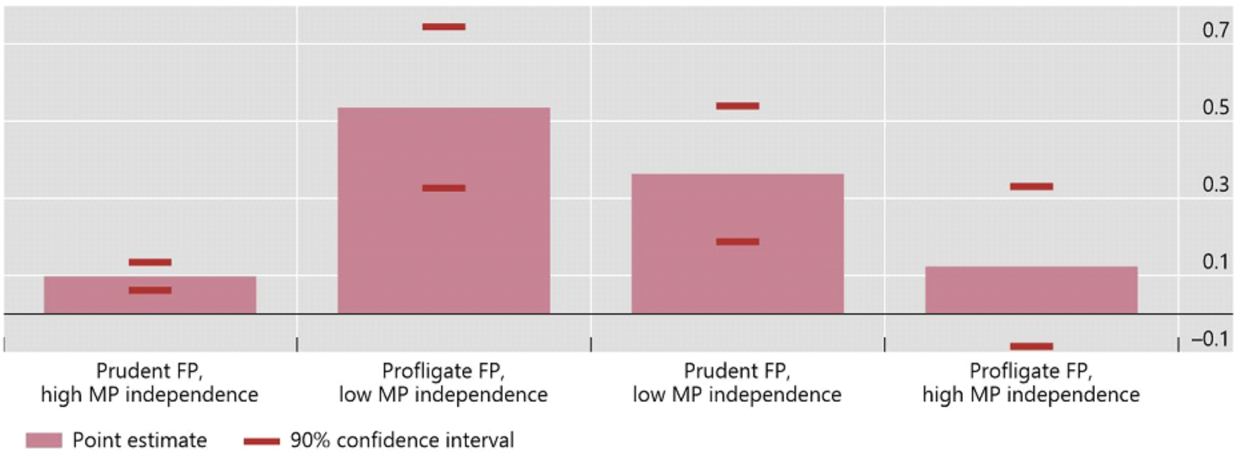


Figure 1: **The inflationary impact of fiscal stimulus across fiscal and monetary regimes**. The figure shows the estimated average impact of a one percentage point increase in the fiscal deficit on inflation over the next two years across different combinations of fiscal and monetary regimes. Fiscal regimes are classified as prudent or profligate based on Mauro et al. (2015). Monetary regimes are defined as being high or low independence based on legal limitations on central bank lending to the public sector in Romelli (2022).

.

(see for example Taylor (1999), Clarida et al. (2000) and Davig and Leeper (2007)).

Our results show that the regimes matter. Figure 1 shows that the estimated average effect of an increase in the overall fiscal deficit on inflation differs strongly across the four combinations of fiscal and monetary policy regimes that we consider. The lowest inflationary effect is found in regimes with a prudent fiscal authority that stabilises debt levels together with an independent central bank with strong legal limitations that prevent lending to the public sector. In this regime which we define as "monetary led", a one percentage point increase in the fiscal deficit results in a 10 basis point increase in the average inflation rate over the next two years.

By contrast, the greatest inflationary effect is found in the regime which we define as "fiscally led", i.e. in which fiscal policy is profligate and the central bank faces limited constraints on lending to the public sector. Under a fiscally-led regime, a one percentage point increase in the overall deficit raises the inflation rate on average by around 50 basis points, over five times higher than in the monetary-led regime.

The average inflationary effects of higher deficits in the other two regimes lie somewhere between the monetary-led and the fiscally-led regimes. The effect of deficits in the prudent fiscal policy and low monetary independence regime is higher than in the profligate fiscal regime with high monetary policy independence (where the effect is, moreover, not statistically significant). This latter result is perhaps surprising. As shown in Sargent and Wallace (1981) or Leeper (1991), any attempt by a central bank to control inflation when fiscal policy is non-Ricardian leads to explosive inflationary dynamics. However, unlike in many theoretical models, the regimes in our empirical analysis are not fixed or permanent.

Beyond analysing the average effect, we further examine how fiscal and monetary policy regimes affect inflation risks across the entire inflation forecast distribution using the inflation-at-risk methodology developed for panels in Banerjee et al. (2020). We find significant non-linearities with particularly large upside inflation risks following an increase in fiscal deficits. The upside inflation risks, as well as the overall variance of inflation, are considerably higher in the fiscally-led regime compared to the monetary-led regime.

The differences in inflation behaviour in the fiscally-led and monetary-led regimes go beyond the relationship between deficits and inflation. In the fiscally-led regime, the average sensitivity of inflation to output growth is around three times larger than in the monetary-led regime. The sensitivity is even higher in the upper tails of the inflation distribution.

Finally, we use our model to shed light on the seemingly unexpected burst of inflation across advanced economies following the Covid-19 pandemic. The large fiscal and monetary policy stimulus that took place in 2020 may have occurred against the backdrop of laxer fiscal and monetary policy regimes, which are more tolerant of higher and rising public debt and positive deviations of inflation from target, respectively. Given the size of the fiscal stimulus and other macroeconomic variables observed in 2020, forecasts from our model suggest that the high inflation outcomes in 2021 and 2022 appear more consistent with a fiscally-led regime rather than a monetary-led regime.

Our paper is related to several streams of research. It adds to the literature examining how inflation depends on the interactions and policy priorities of fiscal and monetary policy makers (e.g. Sargent and Wallace (1981), Leeper (1991), Leeper et al. (2017)). For the United States, Bianchi and Ilut (2017) show that monetary policy accommodation of fiscal policy during the 1960s and 1970s was an important driver of high inflation. Moreover, they find that tight monetary policy on its own was not sufficient to stabilise inflation, noting that inflation in the United States dropped only when agents' beliefs changed about the government's desire to stabilise debt. Bianchi and Melosi (2022) and Bianchi et al. (2022) show that movements in trend inflation in the United States can be accounted for by fiscal shocks and changes in the fiscal-monetary policy mix. We contribute to this literature by empirically classifying the different fiscal and monetary policy regimes and then examining how their interaction has influenced inflation rates in historical cross-country data.

Similarly underscoring the importance of policy interaction, the recent literature on fiscal multipliers shows that the strength of the macroeconomic effects of fiscal policy crucially depends on how a fiscal expansion is financed and how monetary policy responds (see e.g. Woodford (2011), Erceg and Lindé (2014), Ramey (2019), Ascari et al. (2023)). Relative to this literature, we examine how the fiscal and monetary regimes influence the effects of larger fiscal deficits.

Our paper is also related to earlier research by Catao and Terrones (2005) who find a significant link between persistent fiscal deficits and inflation among developing and emerging market economies but a weak or an insignificant relationship for advanced economies. In contrast to this study, we focus on the short-term inflationary impact and on its crucial dependence on the fiscal-monetary policy regime.

Finally, our paper is related to López-Salido and Loria (2020) who find evidence of a structural shift in the dynamics of US inflation risks using an inflation-at-risk framework. We add to this literature by showing how fiscal and monetary regimes have been a source of changing inflation risks over time.

The remainder of the paper is structured as follows. Section 2 describes the estimation methodology, the classification of regimes and the data. Section 3 presents our baseline results, and Section 4 robustness tests and extensions. Finally, Section 5 uses our estimated models to examine how fiscal and monetary policy regimes may have contributed to the burst of inflation in 2021 and 2022. Section 6 concludes.

2 Methodology

We examine the effects of fiscal deficits on inflation by estimating Phillips curve-type models augmented with fiscal deficits, using panel data. We estimate simple linear models as well as quantile regressions that allow us to evaluate inflation tail risks. The models are estimated conditional on four policy regimes that feature different combinations of fiscal and monetary policy. In this section, we first describe the estimated models, then the construction of the four policy regimes and, finally, the data.

2.1 Econometric approach

Our baseline specification to evaluate the effects of deficits on future inflation is as follows:

$$\overline{\pi}_{i,t+1,t+2} = a_i + X'_{it}\beta + \epsilon_{it}. \tag{1}$$

where the dependent variable $\overline{\pi}_{i,t+1,t+2}$ is a simple average of one- and two-year-ahead headline inflation in country i. α_i denotes country fixed effects and $X_{i,t}$ is a vector of explanatory and control variables:

$$X'_{it} = (\Delta def_{it}, \pi_{it}, \Delta y_{it}, \Delta exc_{it}, \Delta oil_{it}). \tag{2}$$

The main covariate of interest is Δdef_{it}, which represents the year-on-year change in fiscal deficit as a percentage of GDP. π_{it} is the current level of headline inflation, year-on-year;[1] Δy_{it} denotes the year-on-year log change in real GDP; Δexc_{it} is the log change in the nominal effective exchange rate, with an increase in exc_{it} denoting an appreciation; and Δoil_{it} denotes the log change in oil prices denominated in local currency. All variables in log changes are expressed as percentages. The model is estimated using ordinary least squares.

In order to examine the possibility that changes in fiscal deficits lead to greater tail risks of inflation, we use novel methods for panel quantile regressions with fixed effects (see Machado and Santos Silva (2019)).We begin with the following location-scale model:

$$\overline{\pi}_{i,t+1,t+2} = \alpha_i + X'_{it}\beta + (\delta_i + X'_{it}\gamma)U_{it}, \tag{3}$$

where $\overline{\pi}_{i,t+1,t+2}$ and $X_{i,t}$ are defined as before. In this model, the size of the coefficients is allowed to vary according to the dependent variable's placement in the conditional inflation distribution. These non-linearities are driven by the scaling of the error term U by a vector of constants γ. The parameters α_i and δ_i denote country i fixed effects. α_i is the time-invariant average level of inflation

[1]While we use the headline inflation rate, we also checked that our results hold if we compute the inflation rate as the log change in the CPI index.

within country i. δ_i is a country-specific, time-invariant scaling parameter of the distribution of U, which has the same properties for all i and t. From Eq (3), we have $\Pr[\delta_i + X_{it}'\gamma > 0] = 1$. The sequence $\{X_{it}\}$ is assumed to be strictly exogenous.[2] U_{it} are unobserved random variables, $i.i.d.$ across countries i and years t, orthogonal to X_{it} and normalised to satisfy $E[U] = 0$ and $E[|U|] = 1$.

We obtain the conditional quantiles for inflation over the next two years using:

$$Q_\pi(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + X_{it}'\beta + X_{it}'\gamma q(\tau), \tag{4}$$

where the scalar $\alpha_i(\tau) = \alpha_i + \delta_i q(\tau)$ is the quantile-τ fixed effect for economy i. $q(\tau)$ denotes the τth quantile of the distribution of the error term U, conditional on X. $\alpha_i(\tau)$ captures the time-invariant effect of individual country characteristics, which potentially vary depending on where the country lies in the conditional inflation distribution. Using this model, we estimate $\beta(\tau) = \beta + \gamma q(\tau)$, for 5 quantiles: 5%, 25%, 50%, 75% and 95%. The confidence intervals are estimated by using a block bootstrapping with 1,000 replications, clustering on country.

For a given country and year, each predicted quantile from Eq (4) represents a point in the CDF $F(\cdot)$ of the inflation forecast. To address noise in our quantile estimates, following Adrian et al. (2019), we interpolate semiparametrically the predicted quantiles using the skewed $t-$distribution (see Azzalini and Capitanio (2003)). The distribution is described by the following function:

$$f(\pi;\mu,\sigma,\alpha,\nu) = \frac{2}{\sigma}t\Big(\frac{\pi-\mu}{\sigma;\nu}\Big)T\left(\alpha\frac{\mu-\pi}{\sigma}\sqrt{\frac{\nu+1}{\nu+\frac{(\pi-\mu)^2}{\sigma}}};\nu+1\right). \tag{5}$$

In Eq (5), $t(\cdot)$ and $T(\cdot)$ are the PDF and the CDF of the distribution, respectively. The distributional parameters μ (location), σ (scale), ν (kurtosis), and α (skewness) are estimated for each country-year pair by minimising the mean squared error between the five predicted quantiles and the distribution-implied values. In other words, we select parameter estimates that minimise the following objective function:

$$(\hat{\mu}_{it+h}, \hat{\sigma}_{it+h}, \hat{\alpha}_{it+h}, \hat{\nu}_{it+h}) = \text{argmin} \sum_{\tau} \big(\hat{Q}_{\pi_{t+h|x_t}}(\tau|x_t) - F^{-1}\big(\tau;\mu,\sigma,\alpha,\nu\big)\big)^2. \tag{6}$$

2.2 Fiscal and monetary policy regimes

We estimate the models – both the linear model and the one for inflation-at-risk – separately for different combinations of fiscal and monetary policy. In particular, we distinguish between four possible policy combinations: “prudent" or “profligate" fiscal policy combined with “high" or “low" monetary policy independence.

[2]In Appendix B we conduct simulation exercises to assess the sensitivity of our estimates to deviations from these key assumptions. We find that such deviations lead our quantile estimates to underestimate the degree of non-linearities present in the data generating process.

2.2.1 Defining the fiscal regime

For the baseline specification, the fiscal regimes are based on a *de facto* measure. We use estimates of a fiscal policy reaction function that considers the response of primary surpluses to the debt-to-GDP ratio (see Bohn (1998)). A positive response is a sufficient condition for fiscal policy to be sustainable, in the sense that the government's intertemporal budget constraint holds. For a given country, the fiscal reaction function is:

$$s_t = \rho d_t + \alpha Z_t + \varepsilon_t, \tag{7}$$

where s_t denotes the primary surplus and d_t denotes the level of government debt-to-GDP at the beginning of the period. Z_t is a vector of control variables that affect the primary balance, such as the business cycle, the transitory component of government spending and commodity prices.

In the baseline model, we use the estimates in Mauro et al. (2015) who operationalise the approach in Bohn (1998). In particular, Mauro et al. (2015) estimate fiscal reaction functions for a panel of countries based on 25-year rolling regressions. Periods of prudent (profligate) fiscal policy are then defined as those with $\rho > 0$ ($\rho < 0$), on condition of a statistically significant coefficient at a minimum of 5% level.

In addition to providing a sufficient statistic on whether the government's intertemporal budget constraint holds, the approach has a number of advantages. Government debt-to-GDP ratios are affected by temporary shocks, such as wartime spending or business cycle fluctuations, which make it difficult to detect violations of the intertemporal budget constraint based on developments of debt alone (see Bohn (1998)).[3] Moreover, the approach is robust to changes in growth rates (g) and interest rates (r), and different debt management policies. However, the relationship between growth rates, interest rates and ρ matter for the trajectory of debt. In particular, if $\rho > (r - g/1 + r)$, the debt ratio is stationary and returns to its initial level after a shock. We also note that considering a regression of the type in Eq (7) implies close correspondence with structural models that feature different types of behaviour of the fiscal authority. For example, in Bianchi and Ilut (2017), passive (active) fiscal policy is defined as one where the fiscal authority is (not) committed to stabilising debt by adjusting taxes.

As noted by Mauro et al. (2015), the approach in Bohn (1998) was developed against the backdrop of rising debt. If the debt ratio is declining, a statistical rejection of a positive ρ would indicate that the intertemporal budget constraint is violated, but in a sense of over-accumulating public assets rather than incurring excessive liabilities. That said, over our sample period, debt ratios were mostly on the rise. In particular, considering all 25-year changes in debt ratios, increasing debt ratios were four times more frequent than decreasing ones, with rising debt ratios over two times larger in absolute value than decreasing ones. (The average 25-year changes are 38 and 18 percentage points, respectively.)

As a robustness test, we also consider *de jure* classifications of fiscal regimes based on whether the economy has in place a fiscal rule for the budget balance - that is, legal numerical limits on the overall balance, the structural or cyclically adjusted balance, or the balance over the cycle.

[3] Indeed, as we discuss in the next section, in the monetary-led regime where fiscal policy is sustainable, debt levels are on average higher than in the fiscally-led regime.

2.2.2 Defining the monetary regime

As to monetary policy, the baseline regimes are based on *de jure* indicators of central bank independence. Given the particular importance of monetary accommodation of fiscal policy, our main measure is based on specific limitations placed on central bank lending to the public sector and enshrined in central bank laws.[4] Grilli et al. (1991) note that if the government is able to influence the quantity and conditions on which it borrows from the monetary authority, it affects the creation of base money and decreases the central bank's economic independence. Similarly, Cukierman et al. (1992) consider a central bank with tighter restrictions on lending to the public sector to be more independent in the pursuit of the price stability objective. The authors argue that providing credit to the government would likely be an important channel behind the relationship between the lack of central bank independence and inflation.

The specific measures on central bank independence we use come from Romelli (2022) who follows the entire set of legislative changes to laws concerning the central bank over time. As indices proposed in earlier literature were generally computed at specific points in time, they do not capture the full set of reforms. For a given country-year observation, we classify monetary policy independence to be low (high) if the indicator is below (above) the median of all country-year observations during the sample period. [ADD MORE DETAILS ON THE SPECIFIC INDICES THAT WE USE - RESTRICTING LENDING TO THE GOVERNMENT]

In our case, given that we are interested in the effect of deficits on inflation, considering an exogenous indicator such as central bank independence to classify the monetary regime is arguably preferable to more endogenous alternatives, such as using the *de facto* degree of inflation stabilisation, for example. That said, in robustness tests, we use an alternative *de facto* measure of the monetary policy regime, capturing the degree to which monetary policy acts to stabilise inflation. In particular, we define the regime based on whether the central bank's policy interest rate is below that suggested by a Taylor rule (Taylor (1993)).[5] We also note that the measures of *de jure* central bank independence in Romelli (2022) feature an interesting dynamic relationship with inflation outcomes, such that reforms to central bank legislation tend to follow periods of high inflation.

2.3 Data

The data are annual and cover 21 advanced economies from 1972 onwards – corresponding to the start date of the central bank independence indices.[6] The end date varies by the measures of fiscal

[4]In addition to capturing restrictions on central bank purchases of government debt securities in the primary market, the indicator covers a number of other dimensions, such as whether strict amounts on loans exist and whether the loan terms are controlled by the central bank; whether the borrower can only be government or also other public sector institutions such as state-owned enterprises; whether interest rates are market-determined; and whether the maturities are limited and clearly specified in the central bank legislation.

[5]Still an alternative approach would be to consider *de facto* measures of central bank independence, such as political pressures on central banks recently published in Binder (2021). However, available data for long time periods are, to our knowledge, sparse.

[6]The economies included are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom

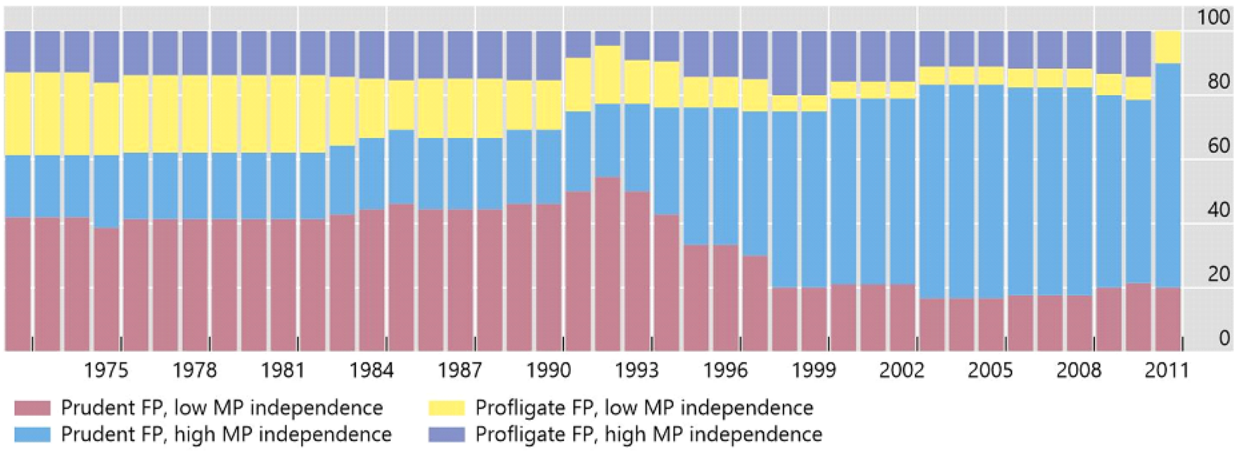


Figure 2: **Fiscal and monetary regimes over time.** The figure shows the share of economies in the four different fiscal and monetary regime combinations. Fiscal regime classification based on Mauro et al. (2015). Prudent FP: Prudent fiscal policy regime, defined as fiscal policy where the primary balance is increasing in the level of debt. Profligate FP: profligate fiscal policy regime, economies where the primary balance is not increasing in the level of debt. Monetary policy regime based on Romelli (2022). High MP independence: high monetary policy independence, defined as central banks with above median *de jure* limitations on lending to the public sector. Low MP independence: low monetary policy independence, defined as central banks with below median *de jure* limitations on lending to the public sector.

and monetary policy regimes used. For the baseline regressions where the fiscal regimes are defined based on Mauro et al. (2015) and monetary policy regimes on Romelli (2022), the sample ends in 2011, covering four decades with diverse monetary and fiscal policy behaviour. The fiscal data are from Mauro et al. (2015). They are extended to more recent periods using data from the IMF Fiscal Monitor. Real GDP and inflation are from national sources and the nominal effective exchange rates are from the BIS. For commodity prices, we use the UNCTAD's commodity price index and the price of West Texas Intermediate (WTI) for oil. The short-term interest rates are from Jordà et al. (2017) and are supplemented by data from the OECD, Datastream and the Global Financial Data database.

3 Baseline results

3.1 Policy regimes over time

The share of economies in the fiscally-led and monetary-led regimes has changed notably over time. Figure 2 shows the share of economies in the different regimes, for each year of the sample. By the early 2010s, around 60-70% of countries were in the monetary-led regime. By contrast, early in the sample, around 20% of economies were in this regime, which included Canada, Germany and the United States. As for the fiscally-led regime, the share was around 25% in the 1970s, dropping to 10% by the 1990s.

and the United States.

The most common combination of regimes early in the sample was the intermediate one featuring prudent fiscal policy and low monetary policy independence. This comprised a number of European countries, together with Japan and New Zealand. Overall, regimes with profligate fiscal policies have been less frequent and their shares have declined further over time.

Perhaps not surprisingly, inflation rates have on average been much higher in regimes with low monetary policy independence (7.7%) than with high independence (4.4%). By contrast, the average inflation rates in profligate and prudent fiscal regimes have been similar, 6.6% in the former and 6.1% in the latter.

While primary deficits have been smaller in prudent than in profligate fiscal regimes (primary deficits of 0.4% vs 1.6% of GDP), overall deficits have been broadly similar in the two regimes (3.0% vs 2.9% of GDP). Moreover, government debt levels have been higher in prudent regimes (55% vs 49% of GDP). Thus, primary fiscal accounts have been closer to balance in economies where overall fiscal deficits have been larger and debt ratios have been higher.

3.2 Average effect of deficits on inflation

Simple least squares estimates show that the relationships between higher deficits and future inflation vary notably between the fiscally-led and monetary-led regimes (see Table 1). The effect is found to be much weaker in the monetary-led regime than in the fiscally-led regime. In the former, a one-percentage-point increase in fiscal deficits is associated with around 0.10 percentage point increase in average inflation over the next two years (first column). By contrast, under a fiscally-led regime, the corresponding effect is over five times as high in magnitude (second column). The effects in both regimes are statistically significant at the 1% level.[7]

In the "intermediate" regimes the effects of deficits on inflation fall in between the two previous ones: point estimates of 0.36 (prudent fiscal and low monetary policy independence; third column) and 0.12 (profligate fiscal and high monetary policy independence; fourth column). However, the latter estimate is not statistically significant at conventional levels.

We also find some relevant results regarding the other control variables in Table 1. Real GDP growth is associated with economically stronger effects on inflation in the fiscally-led regime than in the monetary-led regime. Similar to fiscal deficits, the coefficient on real GDP growth in the intermediate regimes falls between those estimated in the fiscally-led and monetary-led regimes. In all regimes, the coefficient on real GDP growth is statistically significant at the 1% level. Moreover, an exchange rate appreciation obtains the expected negative sign with a statistically significant coefficient only in the monetary-led regime, such that an appreciation is associated with lower future inflation. The statistical insignificance of oil prices - and the negative coefficients in two regimes - owes to the high-frequency fluctuation that is characteristic of commodity prices. Indeed, if we replace future inflation by current inflation as the dependent variable in the estimation, the change in oil prices obtains a statistically significant positive coefficient in all four regimes.

[7] All regressions also include a dummy variable (not shown) that obtains a value of 1 if, for a given country-year observation within a fiscal regime, the same observation is also classified as being in the opposite fiscal regime in another partly overlapping rolling regression in Mauro et al. (2015). In other cases, the dummy variable is assigned a value of zero.

	Monetary-led regime	Fiscally-led regime	Prud FP low MP indep	Profl FP high MP indep
	$\overline{\pi}_{i,t+1,t+2}$	$\overline{\pi}_{i,t+1,t+2}$	$\overline{\pi}_{i,t+1,t+2}$	$\overline{\pi}_{i,t+1,t+2}$
Δdef_{it}	0.0974***	0.536***	0.363***	0.123
	(0.0222)	(0.128)	(0.107)	(0.126)
π_{it}	0.714***	0.722***	0.763***	0.463***
	(0.0263)	(0.0797)	(0.0363)	(0.0628)
Δy_{it}	0.301***	1.005***	0.752***	0.366***
	(0.0390)	(0.101)	(0.102)	(0.0556)
Δexc_{it}	-0.0757**	0.0149	0.0162	-0.0212
	(0.0307)	(0.0271)	(0.0209)	(0.0607)
Δoil_{it}	-0.00109	-0.00811	0.00298	0.00486
	(0.00479)	(0.00559)	(0.00531)	(0.00747)
Observations	314	152	341	126
R-squared	0.747	0.692	0.659	0.391
Number of countries	14	9	13	8

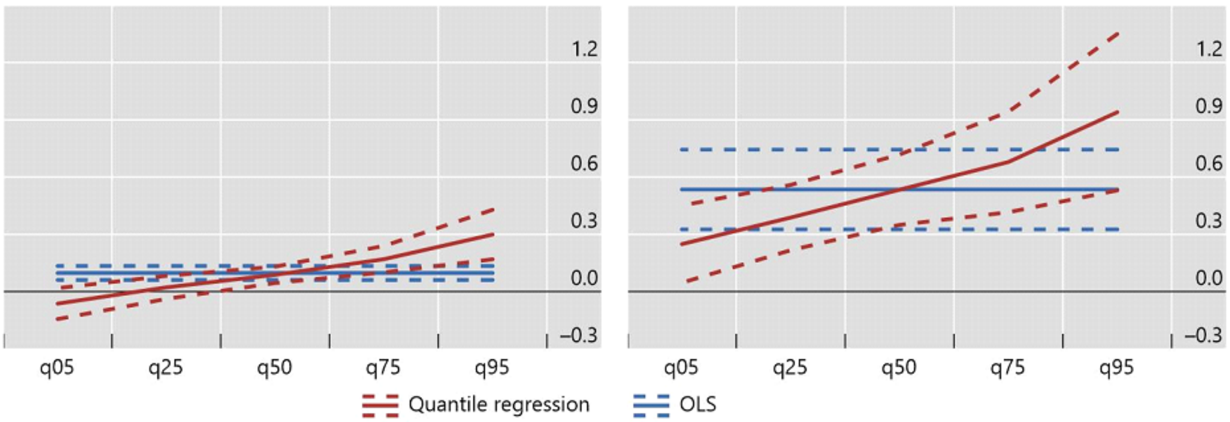
Table 1: **Effects of deficits on inflation across fiscal-monetary regimes, OLS estimates.** This table shows OLS estimates of the relationship between the inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, and changes in fiscal deficit-to-GDP ratio, Δdef_{it}, in year t across different fiscal-monetary policy regimes. Fiscal regimes are classified as prudent (Prud FP) or profligate (Profl FP), based on Mauro et al. (2015). The monetary regimes, low MP indep and high MP indep, are defined as being low or high independence based on the degree of legal limitations on central bank lending to the public sector in Romelli (2022). A monetary-led regime is defined as the combination of prudent fiscal policy and high monetary independence. A fiscally-led regime is defined as profligate fiscal policy and low monetary independence. The control variables are π_{it}: annual inflation rate; Δy_{it}: GDP growth; Δexc_{it}: log change in the nominal effective exchange rate; Δoil_{it}: log change in the local price of oil. The regression also includes country fixed effects. Standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.3 Inflation-at-risk from higher fiscal deficits

The estimates in the previous section showed how future inflation moves, on average, in response to fiscal deficits in the different policy regimes. At the same time, policymakers may want to avoid extreme inflation outcomes and take actions that reduce their likelihood. In this section, we examine the behaviour of the entire inflation forecast distribution under the different policy regimes, as well as the association between higher deficits and tail risks to inflation. We focus on the monetary-led and fiscally-led regimes.

Using the methodology described in Section 2, Figure 3 shows the coefficient on fiscal deficits from the quantile regression, for the monetary-led (left panel) and fiscally-led (right panel) regimes. Moving from left to right within the panels implies moving from lower to higher quantiles, i.e. from the 5% to the 95% quantile.

Figure 3 suggests that higher fiscal deficits increase upside risks to inflation, as the coefficient on deficits is higher in the upper quantiles of the inflation forecast distribution. Moreover, the effect is particularly pronounced in the fiscally-led regime (right panel). In this case, at the 95% quantile, a one percentage point rise in fiscal deficits is associated with close to one percentage point increase in future inflation (see also the last column of Table 2). This effect is around twice



(a) Monetary-led regime

(b) Fiscally-led regime

Figure 3: **Quantile regression estimates of fiscal deficits on inflation.** This figure shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on changes in the fiscal deficit-to-GDP ratio in year t. Coefficients are shown by the q% quantile (x-axis); e.g. q50 denotes the 50% quantile. The left-hand panel shows coefficients estimated in the monetary-led regime while the right-hand panel shows the coefficients estimated in the fiscally-led regime. Quantile estimates are shown with 90% confidence bands using a block bootstrap clustered by country. OLS estimates are shown with 90% confidence bands clustered by country.

as high as at the median and four times as high as at the left tail (5%) of the distribution.

Under a monetary-led regime, deficits also raise upside inflation risks, but the effects are less statistically significant and much lower in economic terms. In this case, the right-tail (95% quantile) shifts by 0.3 ppts when deficits rise by one percentage point (see also the last column of Table 3). Moreover, the effects at the 5% and the 25% quantiles are not statistically different from zero.

The differences in inflation behaviour in the fiscally-led and monetary-led regimes go beyond the relationship between deficits and inflation. In particular, Figure 4 highlights key differences in terms of two moments of the inflation forecast distribution. Setting all variables at their regime-dependent means, the grey distributions show that inflation is higher on average and its variance is larger in the fiscally-led regime than in the monetary-led regime. The red lines show the conditional distributions evaluated at a two standard deviation increase in the change in the fiscal deficit. The conditional distributions shift much further to the right in the fiscally-led regime than in the monetary-led regime.

Fiscally-led and monetary-led regimes also feature differences in terms of sensitivities to real GDP growth (see the third rows in Table 2 and Table 3). In particular, real GDP growth has a stronger relationship with future inflation in the fiscally-led regime than in the monetary-led regime across the entire distribution. At the median of the distribution, the relationship in the fiscally-led regime is over three times as strong as in the monetary-led regime; at the 95% quantile, it is almost four times as strong. Across both distributions, all coefficients on real GDP growth are statistically significant at the 1% level.

At the same time, both fiscally-led and monetary-led regimes display similar non-linearities

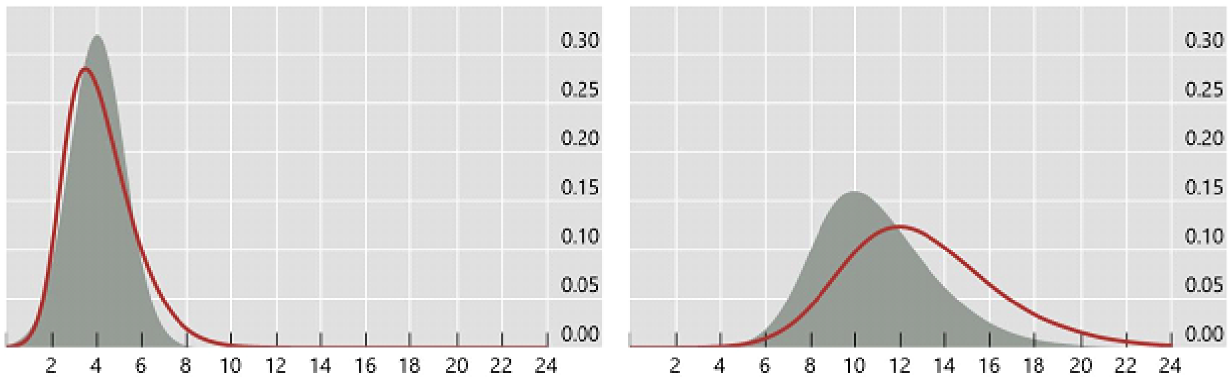
Inflation forecast quantiles	5% $\overline{\pi}_{i,t+1,t+2}$	25% $\overline{\pi}_{i,t+1,t+2}$	50% $\overline{\pi}_{i,t+1,t+2}$	75% $\overline{\pi}_{i,t+1,t+2}$	95% $\overline{\pi}_{t+1,t+2}$
Δdef_{it}	0.251*	0.390***	0.536***	0.679***	0.941***
	(0.145)	(0.106)	(0.111)	(0.169)	(0.245)
π_{it}	0.653***	0.687***	0.722***	0.757***	0.821***
	(0.0878)	(0.0649)	(0.0718)	(0.0772)	(0.115)
Δy_{it}	0.745***	0.872***	1.005***	1.136***	1.374***
	(0.170)	(0.0983)	(0.0766)	(0.0976)	(0.194)
Δexc_{it}	-0.0483**	-0.0175	0.0149	0.0466*	0.105**
	(0.0234)	(0.0202)	(0.0219)	(0.0281)	(0.0447)
Δoil_{it}	-0.000343	-0.00413	-0.00810*	-0.0120*	-0.0191*
	(0.00527)	(0.00390)	(0.00475)	(0.00649)	(0.0106)
Observations	152	152	152	152	152

Table 2: **Quantile regression estimates, fiscally-led regime.** This table shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on changes in the fiscal deficit-to-GDP ratio in year t, Δdef_{it}, annual inflation rate π_{it}, GDP growth, Δy_{it}, log change in the nominal effective exchange rate Δexc_{it}, and log change in the local price of oil, Δoil_{it}. Estimated regressions include quantile-τ fixed effect for economy i. Block bootstrap standard errors clustered by country shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

between current and future inflation. In both regimes inflation persistence is stronger at the right than at the left tail of the forecast distribution, as shown by the higher coefficients on current inflation at the right tail. The finding is consistent with prices being adjusted more frequently at high inflation rates (see e.g. Alvarez et al. (2019)).

For the intermediate regimes, the effects of higher deficits on inflation risks are more mixed. The regime featuring prudent fiscal policy and low monetary policy independence displays similar non-linearities to the fiscally-led and the monetary-led regimes, with higher deficits raising upside inflation risks (see Annex Table A.1). By contrast, when fiscal policy is profligate but monetary independence is high, the effects of deficits are not statistically significant at conventional levels (see Annex Table A.2).

Taken together, the results suggest that in both fiscally prudent and profligate environments, higher deficits are associated with lower future inflation if monetary policy is independent rather than non-independent. These findings appear consistent with previous research highlighting the association between higher central bank independence and lower inflation (e.g. Cukierman et al. (1992); Klomp and Haan (2010); Garriga and Rodriguez (2020)). However, our results suggest that it is not only monetary policy but the combination of fiscal-monetary policy regimes that matters for inflation performance. Relatedly, we also highlight significant differences between the monetary-led and fiscally-led regimes in terms of the conditional mean and variance of future inflation. Moreover, while some earlier studies do not find significant effects of deficits on inflation when inflation is low (e.g. Fischer et al. (2002) and Catao and Terrones (2005)), we also report effects at the lower quantiles of the inflation forecast distribution, but note that such effects are regime-dependent.



(a) Monetary-led regime

(b) Fiscally-led regime

Figure 4: **Fiscal deficits increase inflation by more in a fiscally-led regime.** This figure shows the conditional forecast distribution of the inflation rate (annualised) over the next two years. The grey shaded density shows the conditional distribution evaluated at the sample means of all variables. The red density shows the conditional distribution evaluated at a two standard deviation increase in the change in the fiscal deficit, with other control variables at their means. The left-hand panel shows the conditional distributions of inflation in the monetary-led regime. The right-hand panel shows the conditional distributions of inflation in the fiscally-led regime.

4 Extensions and robustness tests

In this section, we consider a number of extensions and robustness tests to the baseline model, focusing on the fiscally-led and the monetary-led regimes. First, we examine to what extent the inclusion of fiscal deficits improves the out-of-sample forecasting performance of the inflation-at-risk model. Second, we change the way fiscal and monetary policy regimes are defined. Third, we evaluate the robustness of the results to excluding the recent period of low inflation. Fourth, we replace changes in fiscal deficits by a measure of fiscal shocks in the model. Finally, we examine asymmetries between increases and decreases in fiscal deficits in terms of their effect on future inflation.

4.1 Forecasting performance

In examining the out-of-sample predictive ability of our model, our focus is on the extent to which the inclusion of deficits in the Phillips curve type model helps to forecast inflation across the quantiles in the fiscally-led and the monetary-led regimes. To this end, we compute the empirical cumulative distribution of the probability integral transform (PIT; see also Adrian et al. (2019)). We check how closely the fraction of outcomes is to the predicted quantile $Q_\pi(\tau|X_{it})$. Observations close to the 45 degree line between τ and the empirical cumulative distribution would suggest a well calibrated model. Sample uncertainty is accounted for by 95% confidence bands around the 45 degree line.

Figure 5 shows that including deficits, shown by the blue line, helps to improve the forecasting properties of the model relative to the model without deficits (red line). This is especially so in the fiscally-led regime. In that regime, the empirical distribution of the model with fiscal deficits

Inflation forecast quantiles	5% $\overline{\pi}_{i,t+1,t+2}$	25% $\overline{\pi}_{i,t+1,t+2}$	50% $\overline{\pi}_{i,t+1,t+2}$	75% $\overline{\pi}_{i,t+1,t+2}$	95% $\overline{\pi}_{t+1,t+2}$
Δdef_{it}	-0.0630	0.0222	0.0879***	0.170***	0.299***
	(0.0522)	(0.0333)	(0.0297)	(0.0400)	(0.0785)
π_{it}	0.549***	0.636***	0.704***	0.788***	0.921***
	(0.0669)	(0.0418)	(0.0452)	(0.0822)	(0.161)
Δy_{it}	0.254***	0.279***	0.298***	0.322***	0.360***
	(0.0607)	(0.0400)	(0.0389)	(0.0577)	(0.104)
Δexc_{it}	-0.0882**	-0.0816**	-0.0764**	-0.0700**	-0.0599
	(0.0390)	(0.0339)	(0.0316)	(0.0351)	(0.0485)
Δoil_{it}	0.00319	0.000913	-0.000842	-0.00303	-0.00649
	(0.00601)	(0.00455)	(0.00423)	(0.00477)	(0.00651)
Observations	314	314	314	314	314

Table 3: **Quantile regression estimates, monetary-led regime.** This table shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on changes in the fiscal deficit-to-GDP ratio in year t, Δdef_{it}, annual inflation rate π_{it}, GDP growth Δy_{it}, log change in the nominal effective exchange rate Δexc_{it}, and log change in the local price of oil, Δoil_{it}. Estimated regressions include quantile-τ fixed effect for economy i. Block bootstrap standard errors clustered by country shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

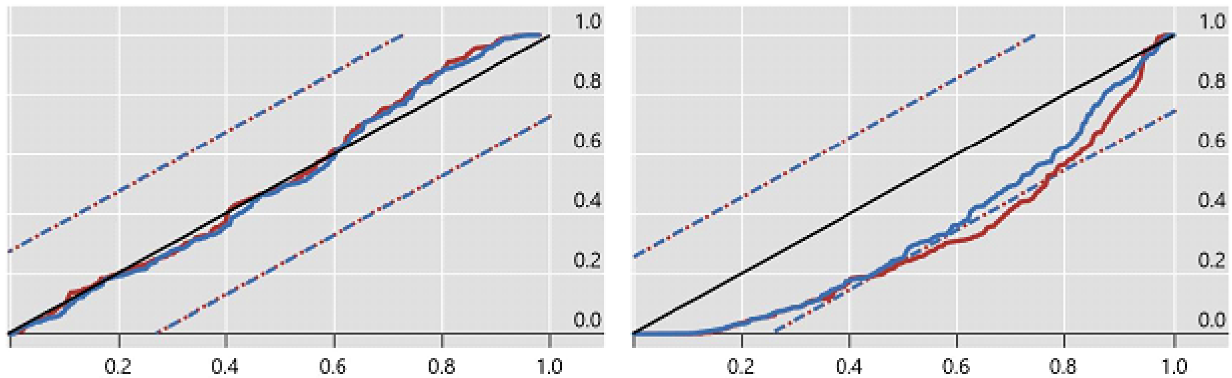
tends to fall closer to the 45 degree line than that of the model without fiscal variables.

4.2 Alternative regime classifications

We then change the classification of the policy regimes. First, instead of using estimates of the fiscal policy reaction function, we define the fiscal regimes based on whether a country has in place a fiscal rule for the budget balance. We draw on the recently published dataset of Davoodi et al. (2022). The data are available from 1985 onwards. We consider rules of both national and supranational types, and covering either the overall balance, the structural or cyclically adjusted balance, or the balance over the cycle. As noted by Schaechter et al. (2012), budget balance rules can help ensure debt sustainability. Moreover, as a *de jure* indicator, the existence of a fiscal rule provides a useful comparison with the *de facto* measure yielded by the fiscal reaction function. Annex Figure A.1 displays the evolution of regimes over time when the presence of fiscal rules is used to define the fiscal regime.

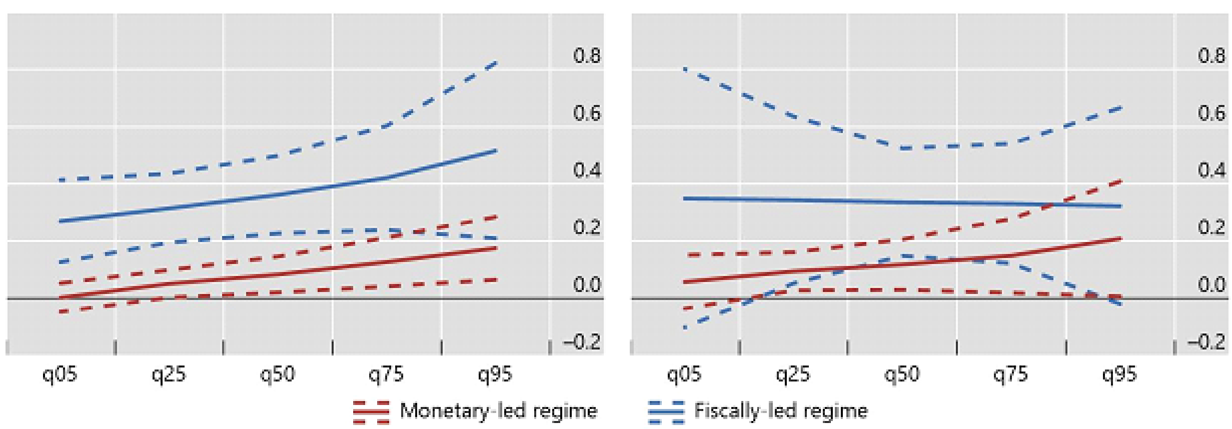
Figure 6, left-hand panel, shows that the results obtained with fiscal rules are similar to the baseline model. In particular, the relationship between deficits and inflation is stronger across the inflation forecast distribution in the fiscally-led regime than in the monetary-led regime. Notably, this result obtains even as the high inflation periods of the 1970s are excluded from the sample due to data availability.[8]

[8]Data on fiscal rules only start in 1985. Given that data for fiscal rules and monetary policy independence are jointly available until 2017, these estimates also cover more years of the low inflation period. In additional robustness tests, we confirm that our results are qualitatively similar if we exclude the high inflation period in our baseline specification.



(a) Monetary-led regime (b) Fiscally-led regime

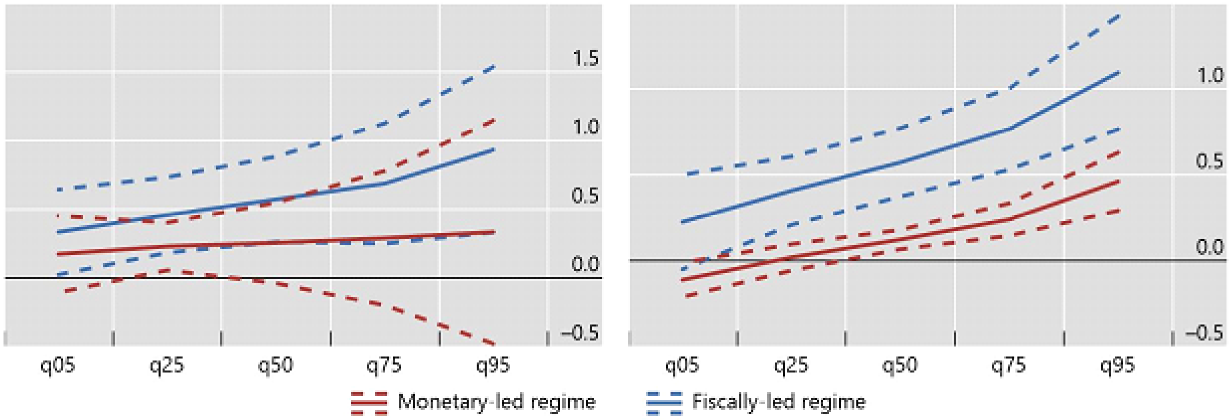
Figure 5: **Cumulative distribution of the probability integral transform in models with and without fiscal deficits.** The x-axis shows the quantile and the y-axis the empirical cumulative distribution. The blue lines show the probability integral transform in the baseline model with changes in fiscal deficits, while the red lines show the probability integral transform in the baseline model without deficits. 95% critical values are included around the 45 degree line.



(a) Fiscal rules (b) Taylor rules

Figure 6: **Quantile regression estimates based on alternative regime classification metrics.** The figure shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on changes in the fiscal deficit-to-GDP ratio in year t. Coefficients are shown by the q% quantile (x-axis); e.g. q50 denotes the 50% quantile. The red lines show estimates in the monetary-led regime while the blue lines show estimates in the fiscally-led regime. The left-hand panel shows estimates when the fiscal regime is based on fiscal rules. Countries with fiscal rules for a budget balance are classified as prudent, those without as profligate. The right-hand panel shows estimates when the monetary regime is based on whether short-term interest rates in the economy are above or below those from estimated Taylor rules. Dotted lines show 90% confidence bands using block bootstraps clustered by country.

Then, we consider a different indicator for the monetary policy regime. We evaluate the extent to which monetary policy has been stabilising, by comparing the level of actual short-term interest rate with that prescribed by a Taylor rule. Periods of interest rates not more than 50 basis



(a) Pre-1994 estimates

(b) Using fiscal shocks

Figure 7: **Quantile regression estimates in pre-1994 sample and fiscal shocks.** The left-hand figure shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on changes in the fiscal deficit-to-GDP ratio in year t estimated over the pre-1994 sample. The right-hand panel shows estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on fiscal shocks measured as deviations from estimated fiscal rules. Coefficients are shown by the q% quantile (x-axis); e.g. q50 denotes the 50% quantile. In both panels, the red lines show estimates in the monetary-led regime while the blue lines show estimates in the fiscally-led regime. Dotted lines show 90% confidence bands using block bootstraps clustered by country.

points below the Taylor prescribed benchmarks are then regarded as stabilising monetary policy. In the opposite case, monetary policy is considered overly accommodative if interest rates are below the Taylor rule benchmark by more than 50 basis points.[9] The Taylor rule parameters are based on Hofmann and Bogdanova (2012).[10] We plot the evolution of the regimes over time in Annex Figure A.2, for the case where the baseline fiscal regimes obtained from fiscal reaction functions are combined with monetary regimes based on Taylor rules.

The right-hand panel of Figure 6 shows that the results are robust to defining monetary regimes with this *de facto* measure based on whether interest rates are above or below those prescribed by our estimated Taylor rules. Consistent with the baseline findings, the effects on inflation from changes in fiscal deficits are higher in economies where interest rates are below those prescribed by our estimated Taylor rules and the fiscal authority is profligate. However, the degree of non-linearity is smaller than in the baseline results for the fiscally-led regime. Moreover, the confidence bands are wide. The estimated coefficients on fiscal deficit changes are lower in regimes where interest rates are above those prescribed by Taylor rules and the fiscal regime is prudent.

[9]The 50 basis points adjustment is done in order to avoid classifying regimes as overly accommodative when their interest rates are close to Taylor rule benchmarks. Moreover, while we use CPI as the relevant price index for all economies, the official inflation target for the US is specified in terms of the PCE index, for which inflation tends to be around 0.5 percentage points below that for the CPI.

[10]See in particular the footnote to Graph 1 in Hofmann and Bogdanova (2012).

4.3 Excluding the low-inflation period

Next, we evaluate the robustness of the results to excluding the recent period of low inflation. The frequency of economies in the monetary-led regime has increased notably over time, while at the same time inflation has trended down in all economies. This raises the question of whether our results for the differences between the fiscally-led and the monetary-led regimes mostly capture this “time effect" of lower inflation that has occurred concurrently with economies shifting from fiscally-led to monetary-led regimes. To examine this issue, we estimate the model for the pre-1994 period that features considerable heterogeneity in terms of the regimes across countries. The left-hand panel of Figure 7 shows that the results are robust to the exclusion of the low inflation period from the sample, but the degree of non-linearity across quantiles is generally smaller.

4.4 Using fiscal shocks

As deficits could be correlated with and partly endogenous to some other explanatory variables, in particular GDP growth, we replace fiscal deficits by a more exogenous measure of fiscal policy. Following the approach of Corsetti et al. (2012) who identify fiscal shocks as residuals from an estimated spending rule, we estimate in a panel set-up a fiscal rule that links primary deficits to lagged primary deficits, the lagged level of government debt and the output gap. Then, we use the residual from this regression as an exogenous measure of fiscal expansion. The right-hand panel of Figure 7 confirms that using expansionary fiscal shocks yields similar results to overall deficits, in particular large differences between the fiscally-led and the monetary-led regime.

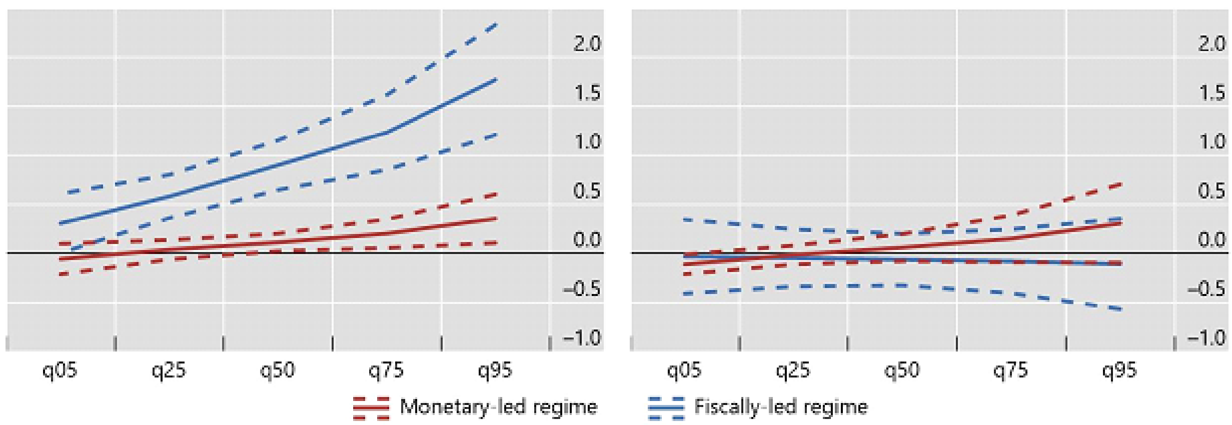
4.5 Examining asymmetries

Next, we examine asymmetries between increases and decreases in fiscal deficits in terms of their effects on future inflation. To do this, we include positive and negative changes in deficits as separate explanatory variables. Figure 8 shows that the effects on inflation stem from increases rather than from decreases in deficits, as the coefficient on the latter is close to zero and statistically insignificant across both fiscally-led and monetary-led regimes.[11]

5 Inflation and Covid-19

In this section, we use our model to shed light on the sudden burst of inflation following the Covid-19 pandemic. The shift in consumption spending from services to goods and supply bottlenecks are important factors that might have contributed to it. At the same time, in many economies the fiscal stimulus has been exceptionally large and monetary policy has remained largely accommodative during the recovery phase (see e.g. BIS (2022)).

[11]Previous literature has highlighted asymmetric effects of contractionary vs expansionary monetary policy on prices, see e.g. Barnichon and Matthes (2018) and Debortoli et al. (2020).



(a) Increases in deficits

(b) Decreases in deficits

Figure 8: **Quantile regression estimates, increases and decreases in fiscal deficits.** The left-hand figure shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on increases in the fiscal deficit-to-GDP ratio in year t while the right-hand panel shows estimates of inflation over the next two years on decreases in the fiscal deficit-to-GDP ratio. Coefficients are shown by the q% quantile (x-axis); e.g. q50 denotes the 50% quantile. In both panels, the red lines show estimates in the monetary-led regime while the blue lines show estimates in the fiscally-led regime. Dotted lines show 90% confidence bands using block bootstraps clustered by country.

Such a strong macroeconomic stimulus followed a period in which the tenets of sound macroeconomic policy had also been questioned. In particular, in the years preceding the pandemic, persistently low inflation and interest rates had strengthened the belief that economies could sustain higher public debt levels and that countries should not rush to reverse fiscal policy lest they jeopardise the recovery. Indeed, many commentators attributed the sluggish growth in the years following the GFC to the rapid reversal of fiscal policy in 2010-11 and warned against making the same mistake in the exit from the pandemic. In addition, with inflation persistently low and nominal policy rates at or close to their effective lower bound pre-pandemic, many central banks judged that downside risks to employment and inflation had increased. In other words, the recent years may potentially represent a shift towards laxer fiscal and monetary policies. Based on our empirical findings, such a shift would imply a stronger impact of fiscal policy on inflation.

To examine this hypothesis, we perform a forecasting exercise using the estimated OLS coefficients for the fiscally-led and monetary-led regimes and data for 2020 as an input. Based on the forecasts shown in Figure 9 the high inflation outcomes following the Covid-19 pandemic appear more consistent with fiscally-led rather than with monetary-led regimes: for three quarters of the sample economies, actual outcomes during 2021-22 fall within the confidence intervals under the fiscally-led regime; by contrast, only for two countries is the inflation outcome consistent with a monetary-led regime. Our results therefore support the hypothesis that the recent burst of inflation also owes to the strong macroeconomic stimulus and a potential change in the regime in which fiscal and monetary policies operate.

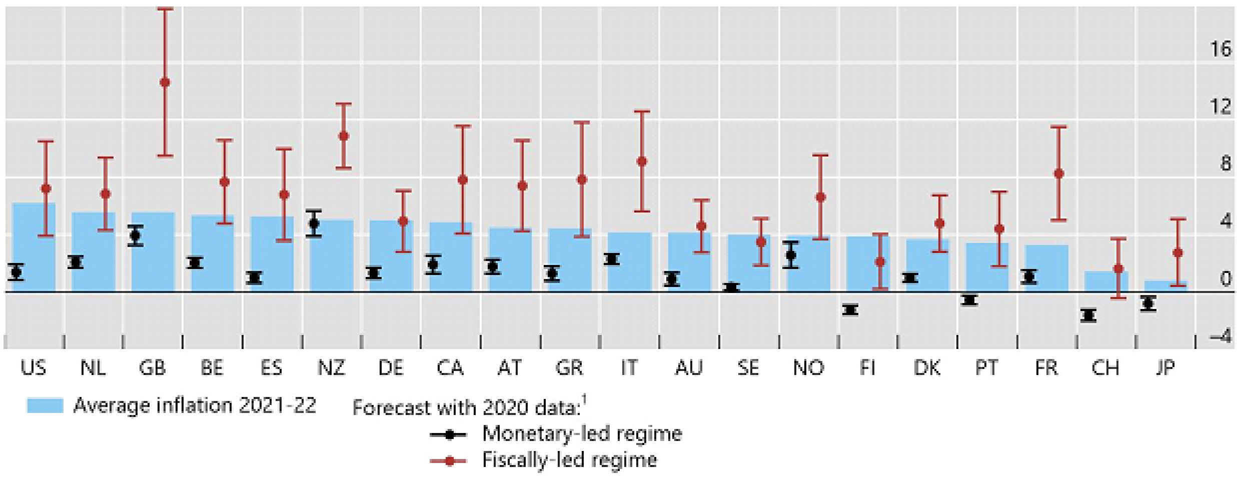


Figure 9: **Inflation outcomes during Covid-19 compared with forecasts under monetary-led and fiscally-led regimes.** The figure shows the average inflation outcomes in 2021-22 (blue bars) and the model-implied forecasts under a monetary-led regime (black line) and a fiscally-led regime (red line). The dot corresponds to the point forecast and the error bars to the 99% confidence interval.

6 Conclusions

Using data for a panel of 21 advanced economies over four decades, this paper shows that the association between higher deficits and future inflation crucially depends on the underlying fiscal and monetary policy regimes. In particular, the inflationary consequences are significantly stronger under a fiscally-led regime, i.e. when the government places less emphasis on stabilising debt and when monetary policy is less committed to price stability. Moreover, both the mean and variance of future inflation are higher under a fiscally-led regime compared to a monetary-led regime. We also show that the relationship between deficits and inflation varies across the conditional inflation distribution, being stronger at the right tail of the distribution, especially under a fiscally-led regime. These results are robust to different approaches of identifying the policy regimes, as well as to excluding the recent period of low inflation from the analysis.

Our findings suggest that changes in policy frameworks may have a sizeable impact on inflation. First, changes in fiscal frameworks, which reduce fiscal discipline or make increasing public debt levels more tolerable, may increase upside inflation risks. Second, recent reviews of monetary policy strategy, such as for example the Federal Reserve’s adoption of average inflation targeting, may have raised the inflationary effect of recent fiscal stimulus. In light of surprisingly high inflation following the Covid-19 pandemic, this may be a fruitful avenue for future research.

Although our findings are based on the long-run historical experience of inflation in advanced economies, our results also have lessons for emerging market and less developed economies. These include the importance of fiscal as well as monetary frameworks, and their interaction in influencing the mean, volatility and upside risks to inflation. Our findings suggest that similar analysis for emerging economies could shed light on the differential success in taming inflation in emerging Asia compared with that in Latin America.

References

Adrian, T., N. Boyarchenko, and D. Giannone (2019): "Vulnerable growth," *American Economic Review*, 109, 1263–1289.

Alvarez, F., M. Beraja, M. Gonzalez-Rozada, and P. A. Neumeyer (2019): "From hyperinflation to stable prices: Argentina's evidence on menu cost models," *The Quarterly Journal of Economics*, 134, 451–505.

Ascari, G., P. Beck-Friis, A. Florio, and A. Gobbi (2023): "Fiscal foresight and the effects of government spending: It's all in the monetary-fiscal mix," *Journal of Monetary Economics*, 134, 1–15.

Azzalini, A. and A. Capitanio (2003): "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution," *Journal of the Royal Statistical Society Series B*, 65, 367–389.

Banerjee, R. N., J. Contreras, A. Mehrotra, and F. Zampolli (2020): "Inflation at risk in advanced and emerging market economies," BIS Working Papers 883, Bank for International Settlements.

Barnichon, R. and C. Matthes (2018): "Functional Approximation of Impulse Responses," *Journal of Monetary Economics*, 99, 41–55.

Bianchi, F., R. Faccini, and L. Melosi (2022): "A Fiscal Theory of Trend Inflation," Working Paper 30727, National Bureau of Economic Research.

Bianchi, F. and C. Ilut (2017): "Monetary/Fiscal policy mix and agents' beliefs," *Review of Economic Dynamics*, 26, 113–139.

Bianchi, F. and L. Melosi (2022): "Inflation as a Fiscal Limit," Working Paper Series WP 2022-37, Federal Reserve Bank of Chicago.

Binder, C. C. (2021): "Political pressure on central banks," *Journal of Money, Credit and Banking*, 53, 715–744.

BIS (2022): *Annual Economic Report 2022*, Bank for International Settlements.

Bohn, H. (1998): "The behavior of US public debt and deficits," *The Quarterly Journal of Economics*, 113, 949–963.

Catao, L. A. and M. E. Terrones (2005): "Fiscal deficits and inflation," *Journal of Monetary Economics*, 52, 529–554.

Clarida, R., J. Galí, and M. Gertler (2000): "Monetary policy rules and macroeconomic stability: evidence and some theory," *The Quarterly Journal of Economics*, 115, 147–180.

CORSETTI, G., A. MEIER, AND G. J. MÜLLER (2012): "What determines government spending multipliers?" *Economic Policy*, 27, 521–565.

CUKIERMAN, A. (1992): "Central bank strategy, credibility, and independence: theory and evidence," *Journal des Économistes et des Études Humaines*, 3, 581–590.

CUKIERMAN, A., S. B. WEB, AND B. NEYAPTI (1992): "Measuring the independence of central banks and its effect on policy outcomes," *The World Bank Economic Review*, 6, 353–398.

DAVIG, T. AND E. M. LEEPER (2007): "Generalizing the Taylor principle," *The American Economic Review*, 97, 607–635.

DAVOODI, H. R., A. FOTIOU, P. ELGER, D. GARCIA-MACIA, X. HAN, A. LAGERBORG, W. R. LAM, AND P. MEDAS (2022): "Fiscal rules and fiscal councils. Recent trends and performance during the COVID-19 pandemic," IMF Working Papers 2022/11, International Monetary Fund.

DEBORTOLI, D., M. FORNI, L. GAMBETTI, AND L. SALA (2020): "Asymmetric Effects of Monetary Policy Easing and Tightening," CEPR Discussion Papers 15005, Centre for Economic Policy Research.

ERCEG, C. AND J. LINDÉ (2014): "Is there a fiscal free lunch in a liquidity trap?" *Journal of the European Economic Association*, 12, 73–107.

FISCHER, S., R. SAHAY, AND C. A. VÉGH (2002): "Modern hyper-and high inflations," *Journal of Economic Literature*, 40, 837–880.

GARRIGA, A. C. AND C. M. RODRIGUEZ (2020): "More effective than we thought: Central bank independence and inflation in developing countries," *Economic Modelling*, 85, 87–105.

GRILLI, V., D. MASCIANDRO, G. TABELLINI, E. MALINVAUD, AND M. PAGANO (1991): "Political and monetary institutions and public financial policies in the industrial countries," *Economic Policy*, 6, 341–392.

HOFMANN, B. AND B. BOGDANOVA (2012): "Taylor rules and monetary policy: a global "Great Deviation"?" *BIS Quarterly Review*, September, 37–49.

JORDÀ, O., M. SCHULARICK, AND A. M. TAYLOR (2017): "Macrofinancial history and the new business cycle facts," *NBER Macroeconomics Annual*, 31, 213–263.

KLOMP, J. AND J. D. HAAN (2010): "Inflation and central bank independence: a meta-regression analysis," *Journal of Economic Surveys*, 24, 593–621.

KRUGMAN, P. (2021): "Fighting Covid is like fighting a war," *The New York Times*.

LANDAU, J.-P. (2021): "Inflation and the Biden stimulus," VoxEU column February 8, 2021.

LEEPER, E. M. (1991): "Equilibria under 'active' and 'passive' monetary and fiscal policies," *Journal of Monetary Economics*, 27, 129–147.

LEEPER, E. M., N. TRAUM, AND T. B. WALKER (2017): "Clearing Up the Fiscal Multiplier Morass," *American Economic Review*, 107, 2409–54.

LÓPEZ-SALIDO, J. D. AND F. LORIA (2020): "Inflation at Risk," Finance and Economics Discussion Series 2020-013, Board of Governors of the Federal Reserve System (U.S.).

MACHADO, J. A. AND J. SANTOS SILVA (2019): "Quantiles via moments," *Journal of Econometrics*, 213, 145–173.

MAURO, P., R. ROMEU, A. BINDER, AND A. ZAMAN (2015): "A modern history of fiscal prudence and profligacy," *Journal of Monetary Economics*, 76, 55–70.

NICKELL, S. J. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 1417–1426.

RAMEY, V. A. (2019): "Ten Years after the Financial Crisis: What Have We Learned from the Renaissance in Fiscal Research?" *Journal of Economic Perspectives*, 33, 89–114.

ROMELLI, D. (2022): "The political economy of reforms in central bank design: evidence from a new dataset," *Economic Policy*, 37, 641–688.

SARGENT, T. J. AND N. WALLACE (1981): "Some unpleasant monetarist arithmetic," *Quarterly Review, Federal Reserve Bank of Minneapolis*, 5.

SCHAECHTER, A., T. KINDA, N. BUDINA, AND A. WEBER (2012): "Fiscal rules in response to the crisis - toward the 'next-generation' rules. A new dataset," IMF Working Papers 2012/187, International Monetary Fund.

SUMMERS, L. (2021): "The Biden stimulus is admirably ambitious. But it brings some big risks, too." *The Washington Post*.

TAYLOR, J. B. (1993): "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

——— (1999): "A Historical analysis of monetary policy rules," in *Monetary Policy Rules*, National Bureau of Economic Research, Inc, NBER Chapters, 319–348.

THE ECONOMIST (2021): "Has the pandemic shown inflation to be a fiscal phenomenon," *www.economist.com/finance-and-economics/2021/12/18/has-the-pandemic-shown-inflation-to-be-a-fiscal-phenomenon*.

WOODFORD, M. (2011): "Simple Analytics of the Government Expenditure Multiplier," *American Economic Journal: Macroeconomics*, 3, 1–35.

Appendix A: Tables and Figures

Inflation forecast quantiles	5% $\overline{\pi}_{t+1,t+2}$	25% $\overline{\pi}_{t+1,t+2}$	50% $\overline{\pi}_{t+1,t+2}$	75% $\overline{\pi}_{t+1,t+2}$	95% $\overline{\pi}_{t+1,t+2}$
Δdef_{it}	0.244***	0.304***	0.350***	0.412***	0.526**
	(0.0914)	(0.0835)	(0.0975)	(0.139)	(0.213)
π_{it}	0.541***	0.652***	0.737***	0.853***	1.066***
	(0.0448)	(0.0281)	(0.0324)	(0.0627)	(0.113)
Δy_{it}	0.681***	0.716***	0.743***	0.780***	0.848***
	(0.133)	(0.101)	(0.0972)	(0.118)	(0.206)
Δexc_{it}	-0.0223	-0.00310	0.0117	0.0317	0.0686
	(0.0241)	(0.0169)	(0.0193)	(0.0305)	(0.0531)
Δoil_{it}	0.00655	0.00478	0.00340	0.00154	-0.00187
	(0.00595)	(0.00422)	(0.00458)	(0.00679)	(0.0125)
Observations	341	341	341	341	341

Table A.1: **Quantile regression estimates in prudent fiscal policy, low independence monetary policy regimes.** This table shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on changes in the fiscal deficit-to-GDP ratio in year t, Δdef_{it}, annual inflation rate π_{it}, GDP growth Δy_{it}, log change in the nominal effective exchange rate Δexc_{it}, and log change in the local price of oil, Δoil_{it}. Estimated regressions include quantile-τ fixed effect for economy i. Block bootstrap standard errors clustered by country shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Inflation forecast quantiles	5% $\overline{\pi}_{t+1,t+2}$	25% $\overline{\pi}_{t+1,t+2}$	50% $\overline{\pi}_{t+1,t+2}$	75% $\overline{\pi}_{t+1,t+2}$	95% $\overline{\pi}_{t+1,t+2}$
Δdef_{it}	0.119	0.121	0.123	0.125	0.127
	(0.203)	(0.101)	(0.120)	(0.204)	(0.333)
π_{it}	0.326***	0.399***	0.454***	0.512***	0.608***
	(0.110)	(0.0862)	(0.0917)	(0.119)	(0.189)
Δy_{it}	-0.00834	0.191***	0.341***	0.501***	0.764***
	(0.183)	(0.0656)	(0.0685)	(0.104)	(0.180)
Δexc_{it}	-0.0427	-0.0312	-0.0226	-0.0134	0.00170
	(0.0526)	(0.0530)	(0.0620)	(0.0790)	(0.104)
Δoil_{it}	0.0137*	0.00900	0.00546	0.00166	-0.00456
	(0.00711)	(0.00645)	(0.00689)	(0.00819)	(0.0110)
Observations	126	126	126	126	126

Table A.2: **Quantile regression estimates in profligate fiscal policy, high independence monetary policy regimes.** This table shows the estimated coefficients in quantile regressions of inflation rate over the next two years (annualised) in country i, $\overline{\pi}_{i,t+1,t+2}$, on changes in the fiscal deficit-to-GDP ratio in year t, Δdef_{it}, annual inflation rate π_{it}, GDP growth Δy_{it}, log change in the nominal effective exchange rate Δexc_{it}, and log change in the local price of oil, Δoil_{it}. Estimated regressions include quantile-τ fixed effect for economy i. Block bootstrap standard errors clustered by country shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

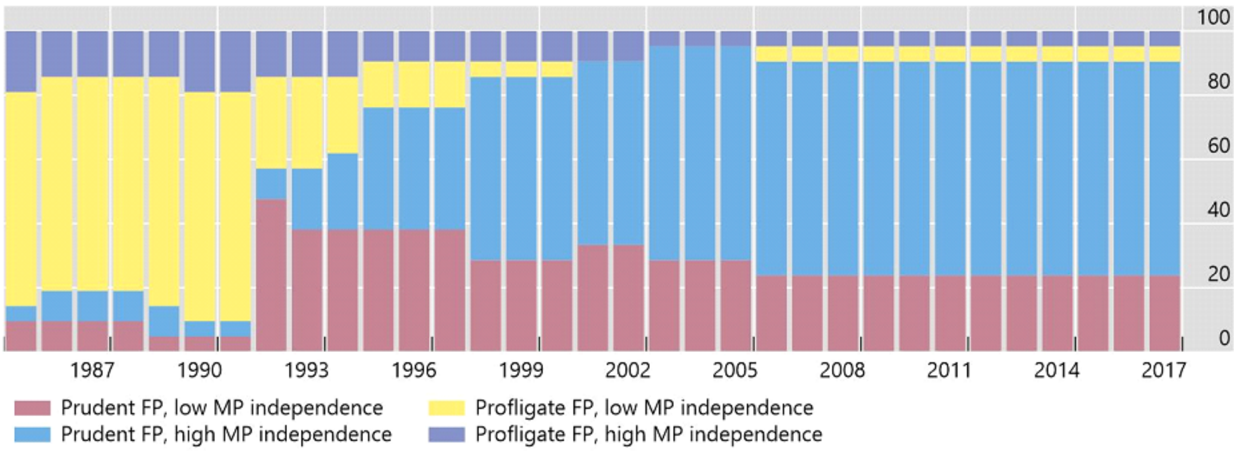


Figure A.1: **Fiscal and monetary regimes over time, based on *de jure* fiscal rules to classify fiscal regimes.** The figure shows the share of economies in the four different fiscal and monetary regime combinations. Fiscal regime classification is based *de jure* fiscal rules. Countries with fiscal rules for a balanced budget are classified as prudent, those without as profligate. Monetary policy regime based on Romelli (2022). High MP independence: high monetary policy independence, defined as central banks with above median *de jure* limitations on lending to the public sector. Low MP independence: low monetary policy independence, defined as central banks with below median *de jure* limitations on lending to the public sector.

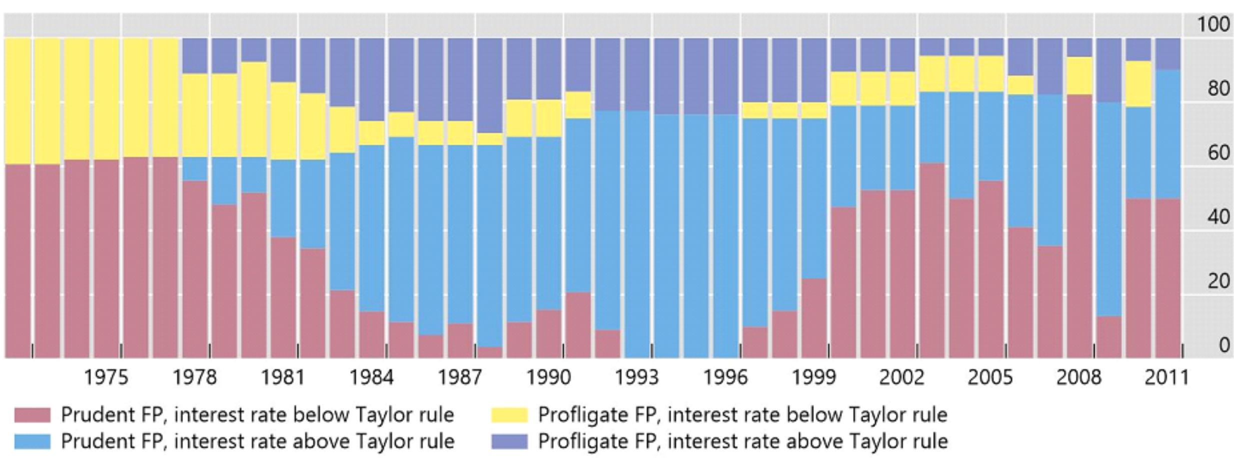


Figure A.2: **Fiscal and monetary regimes over time, based on deviations from Taylor rules to classify monetary regimes.** The figure shows the share of economies in the four different fiscal and monetary regime combinations. Fiscal regime classification based on Mauro et al. (2015). Prudent FP: Prudent fiscal policy regime, defined as fiscal policy where the primary balance is increasing in the level of debt. Profligate FP: profligate fiscal policy regime, economies where the primary balance is not increasing in the level of debt. Monetary policy regime based on whether short-term interest rates in the economy are above or below those from estimated Taylor rules.

Appendix B: Robustness of estimation techniques

The quantiles via moments estimation procedure of Machado and Santos Silva (2019) solves a number of challenges in extending quantile regression methods to panel data, but the asymptotic proofs require certain assumptions about the data generating process (DGP) that may not hold in our data. In this appendix, we examine the sensitivity of estimates to deviations from the key assumption that the sequence $\{X_{it}\}$ of regressors is assumed to be strictly exogenous and $i.i.d.$ for any country i and independent across i. Two factors are likely to lead to deviations from this assumption. First, inflation persistence leads to serial correlation in the errors. As is well known from time-series econometrics, this can lead to a bias in small samples. In addition, in a panel setting with fixed effects, this can lead to an additional source of bias (Nickell (1981)).[12] Second, interconnections across countries, most clearly within the euro area through correlations in the nominal effective exchange rate, but also through other factors such as common oil shocks and global value chains, would violate the assumption of independent regressors across countries.

B.1 Monte Carlo simulation

In the main results of this paper, we document significant non-linearities in the effects of lagged inflation and fiscal deficits across the inflation distribution in advanced economies. We verify the robustness of our estimation technique using a Monte Carlo simulation, in which we explore a few departures of our data from the assumptions used to derive the location-scale model in Machado and Santos Silva (2019). In particular, using a simulated data set, we show that through the effect of noise due to persistence and cross-correlation in the regressors, the quantile regression estimation appears to understate the true degree of non-linearities in the simulated data. In the context of our real-world data, the simulation exercise suggests that the non-linearities in the effects of fiscal deficits and lagged inflation may be even larger than our reported estimates.

We describe the simulation technique and results in further detail below.

For the Monte Carlo exercise we restrict the number of countries and time periods to match our baseline sample of $n = 21$ and $T = 40$. We then simulate time series for our dependent variable inflation and the regressors.

For all countries, regardless of type, we characterize the DGP as follows:

- Each country is assigned two fixed effects, as in Machado and Santos Silva (2019). The first fixed effect, α_i, corresponds to the country-specific time-invariant average inflation. For each country, this fixed effect is drawn randomly from a normal distribution with mean 5 and standard deviation 2.[13] The second fixed effect, δ_i, describes the countries' time invariant average level of scaling applied to the error term. Intuitively, the second fixed effect allows inflation in some countries to respond more or less strongly to random

[12] Machado and Santos Silva (2019) investigate potential bias arising from fixed effects in quantile regressions. They find that the bias is not too large for $n/T < 10$. In our case $n = 21$ and $T = 40$.

[13] These moments were selected based on an approximation of the average inflation distribution across advanced economies. The estimation results are relatively insensitive to the choice of moments.

shocks relative to other countries in the sample. For each country, the second fixed effect is randomly drawn from the standard normal distribution.

- $\pi_{i,t} \sim AR(1), \epsilon_{i,t} \sim \mathcal{N}(\mu_\pi, \sigma_\pi^2)$
- $\Delta def_{i,t} \sim \mathcal{N}(\mu_{\Delta def}, \sigma_{\Delta def}^2)$
- $\Delta y_{i,t} \sim \mathcal{N}(\mu_{\Delta y}, \sigma_{\Delta y}^2)$
- $\Delta oil_t \sim \mathcal{N}(\mu_{\Delta oil}, \sigma_{\Delta oil}^2)$

We assume that inflation $\pi_{i,t}$ is an AR(1) process, while fiscal deficits $\Delta def_{i,t}$, output growth $\Delta y_{i,t}$, and oil shocks $\Delta oil_{i,t}$ are assumed to be $i.i.d.$ with the means and variances taken from the unconditional moments of our data. In addition, we allow for cross-correlation across countries in our simulated exchange rate variable $\Delta exc_{i,t}$. In particular, our simulation assumes two types of countries:

Type 1 (Non-euro area country): Country 1 exchange rate growth is uncorrelated with euro exchange rate growth, i.e., $\Delta exc_{1,t} \sim \mathcal{N}_1(\mu_{\Delta exc}, \sigma_{\Delta exc}^2)$.

Type 2 (Country in monetary union): Exchange rate growth is perfectly correlated for later years of the sample due to the introduction of a common currency.[14] For all t, we assume $\Delta exc_{2,t}$ is drawn from two multinomial distributions,

$$\Delta exc_2 \sim \begin{cases} \mathcal{N}_2(\mu_{\Delta exc}, \Sigma_2), & t < t_{EUR} \\ \mathcal{N}_3(\mu_{\Delta exc}, \Sigma_3), & t \geq t_{EUR} \end{cases}$$

where

$$\Sigma_2 = \sigma_{\Delta exc}^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ . & . & . & . \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\Sigma_3 = \sigma_{\Delta exc}^2 \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ . & . & . & . \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

With these assumptions about the variables, our simulated data is generated with the following location-scale model:

$$\overline{\pi}_{i,t+1,t+2} = a_i + X_{it}'\beta + (\delta_i + X_{it}'\gamma)U_{it}, \quad U_{it} \sim \mathcal{N}(0,1) \tag{B1}$$

$$X_{i,t}' = (\Delta def_{i,t}, \pi_{i,t}, \Delta y_{i,t}, \Delta exc_{i,t}, \Delta oil_t) \tag{B2}$$

[14] In principle, other variables could also be correlated across countries, but we set this aside for simplicity to examine the potential bias stemming from one variable.

As before, conditional quantiles are then given by:

$$Q_\pi(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + X_{it}'\beta + X_{it}'\gamma q(\tau). \tag{B3}$$

Since $U|X \sim \mathcal{N}(0,1)$, the conditional quantile of U is obtained using properties of the standard normal distribution. In particular,

$$Q_\pi(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + X_{it}'\beta + X_{it}'\gamma\Phi^{-1}(\tau) \tag{B4}$$

where $\Phi^{-1}(\cdot)$ denotes the inverse CDF of $U \sim N(0,1)$. We estimate the average effects of the regressors using ordinary least squares (OLS), and subsequently the quantile effects using the method of Machado and Santos Silva (2019).

In our simulation exercise, we are primarily concerned with the degree of bias on parameter estimates of $\Delta def_{i,t}$ as well as biases on $\pi_{i,t}$ due to inflation persistence and on $\Delta exc_{i,t}$ in the presence of a monetary union. Results from the Monte Carlo simulation are shown in Table B.1. The results demonstrate that both the regression quantile and OLS estimates recover the β^*s with reasonable accuracy (Figure B.1). Furthermore, we show that the noise resulting from inflation persistence and cross-correlation in the regressors lead to attenuation towards the average effects. In other words, the noise leads to an underestimation of the true degree of non-linearities in the effect of deficits and lagged inflation on the two-period-ahead average inflation. Applying these findings to our main results, this evidence suggests that our estimates likely underestimate the true extent of non-linearities in the real-world inflation distribution. This is particularly noteworthy for observations with above-median inflation, since it suggests that the risk of further inflation due to fiscal deficits may be even higher than estimated.

	Quantile (τ)	5%	25%	50%	75%	90%	OLS	$\beta^*(\tau)$
100 Reps	$\overline{\beta}_{\Delta def}$	0.24 (0.03,0.45)	0.30 (0.21,0.39)	0.33 (0.21,0.44)	0.35 (0.20,0.51)	0.39 (0.14,0.64)	0.33 (0.33,0.33)	$0.33 + 0.22\Phi^{-1}(\tau)$
	$\overline{\beta}_{\pi}$	0.65 (0.40,0.90)	0.70 (0.59,0.80)	0.72 (0.59,0.85)	0.74 (0.55,0.93)	0.77 (0.47,1.06)	0.72 (0.72,0.72)	$0.72 + 0.12\Phi^{-1}(\tau)$
	$\overline{\beta}_{\Delta exc}$	−0.04 (-0.09,0.01)	−0.03 (-0.05,-0.01)	−0.03 (-0.06,-0.003)	−0.03 (-0.06,0.01)	−0.02 (-0.08,0.04)	−0.03 (-0.03,-0.03)	$-0.03 + 0.025\Phi^{-1}(\tau)$
10000 Reps	$\overline{\beta}_{\Delta def}$	0.17 (-0.13,0.47)	0.28 (0.15,0.40)	0.33 (0.22,0.43)	0.37 (0.25,0.50)	0.43 (0.29,0.62)	0.33 (0.32,0.33)	$0.33 + 0.22\Phi^{-1}(\tau)$
	$\overline{\beta}_{\pi}$	0.59 (0.13,1.05)	0.68 (0.49,0.87)	0.72 (0.57,0.88)	0.76 (0.58,0.95)	0.81 (0.53,1.09)	0.72 (0.72,0.72)	$0.72 + 0.12\Phi^{-1}(\tau)$
	$\overline{\beta}_{\Delta exc}$	−0.05 (-0.13,0.03)	−0.04 (-0.07,-0.004)	−0.03 (-0.06,-0.003)	−0.02 (-0.06,0.01)	−0.01 (-0.06,0.03)	−0.03 (-0.03,-0.03)	$-0.03 + 0.025\Phi^{-1}(\tau)$
		N =840 (T = 40)						

Table B.1: **Monte Carlo simulation results.** In each column, we report average estimates of $\beta(\tau) = \beta + \gamma Q_{\overline{\pi}}(\tau|X)$ for simulations with 100 and 10,000 repetitions, respectively. The 90% confidence intervals, shown in parentheses, are computed using the average point estimate and average standard error from the repetitions in each simulation. The simulated results show that OLS estimates are robust to inflation persistence and cross-correlation in the regressors. Non-linearities are also reflected in the simulation.

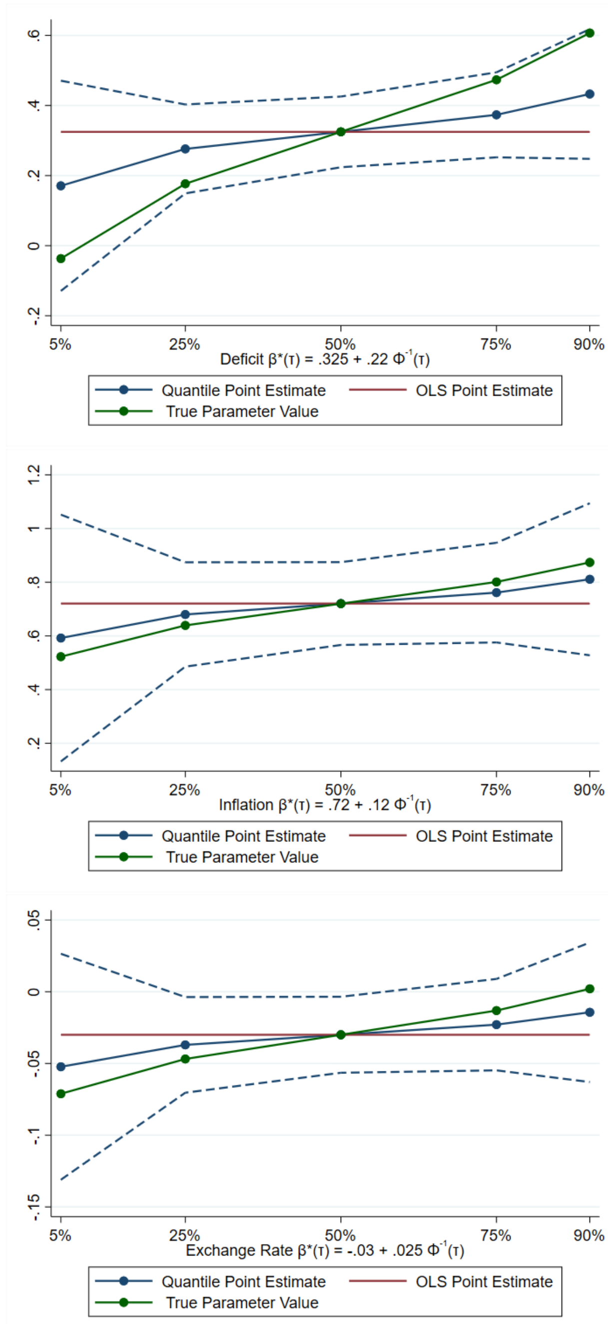


Figure B.1: **Estimated vs. "true" parameter values.** We plot the estimated parameter values against what would be the true parameter values across quantiles based on the DGP described above. The plots suggest that the effects are overstated below the median quantile and understated above it, so the real-world non-linearities are likely understated due to noise in the sample. Importantly, the right tail inflation risk from fiscal deficits may be larger than it seems, based on our simulated results.

Conclusion

This dissertation comprises three chapters at the intersection of household finance and empirical macroeconomics, unified by a central theme: leveraging the applied microeconomics toolkits to enhance our understanding of the intersection of household financial behavior and macroeconomic policy. Each chapter tackles a different domain—credit relief, expectations measurement, and fiscal inflation risk—yet all share a commitment to understanding and improving macroeconomic policy perscriptions based on rigorous findings using rich microeconomic data.

The first chapter demonstrates that well-targeted debt relief programs, such as mortgage forbearance, can yield substantial and lasting improvements in household financial health. Using granular credit data and quasi-experimental variation, the analysis provides rare evidence on the long-run consequences of temporary financial interventions—informing both future crisis response and debates on the optimal design of tageted debt relief and countercyclical stabilization policies.

The second chapter turns to a different but equally consequential channel: expectations. It shows that household inflation expectations—central to many theoretical models and policy discussions—are highly sensitive to the way survey questions are framed. This finding calls for more careful interpretation of expectations data and suggests that macroeconomic forecasting models must account for potential measurement error when relying on household survey responses.

Finally, the third chapter applies microeconometric techniques to a question at the heart of post-pandemic economic policy: when do fiscal deficits become inflationary? By classifying fiscal-monetary regimes and employing an inflation-at-risk framework, the chapter offers a structural interpretation of recent inflation dynamics. It shows that the institutional context in which policy operates—particularly the balance between fiscal and monetary authority—profoundly shapes the macroeconomic risks that follow large-scale stimulus.

Taken together, these essays illustrate how the applied microeconomics toolkit can offer fresh insights into longstanding macroeconomic debates. They underscore the value of combining credible identification strategies, rich data sources, and theoretical awareness to inform real-world policy. As policymakers continue to confront economic uncertainty—whether through designing relief programs, interpreting inflation signals, or navigating the fiscal-monetary interface—the kind of empirical grounding offered in this dissertation becomes increasingly vital.