Title
Essays in Macroeconomics and Comparative Economics

Author
Yaremko, Vitaliia

Publication Date
2023

Peer reviewed|Thesis/dissertation

Essays in Macroeconomics and Comparative Economics

by

Vitaliia Yaremko

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 Yuriy Gorodnichenko, Chair
Professor Gérard Roland
Professor Benjamin Schoefer

Spring 2023

Essays in Macroeconomics and Comparative Economics

Abstract

Essays in Macroeconomics and Comparative Economics

by

Vitaliia Yaremko

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Yuriy Gorodnichenko, Chair

This dissertation studies how institutions, policies, and macroeconomic events shape individual and aggregate outcomes, both from historical and present-day perspectives. The first chapter studies persistent effects of historical institutions on present-day outcomes in the case of a Soviet repressive policy that intended to eradicate market culture and distort social structure in Ukraine. The second chapter examines the interplay between historical and modern institutions and their effect on individual behavior. This chapter studies evidence about the persistence of a socialist-era norm for women to work "double shifts" throughout the market transition period. The third chapter focuses on the effect of current macroeconomic events on individual behavior. Based on novel experimental evidence about the effect of inflation expectations on labor supply, this chapter provides policy-relevant insights about the risk of wage-price spirals in a high inflation setting.

In Chapter 1, I assemble a novel dataset to examine the long-term consequences of blacklisting, a Soviet policy used to deter market-oriented behavior through collective punishment of Ukrainian villages in 1932-33. Under blacklisting, all village residents could be banned from trade and provision of crucial goods, prohibited from moving, and imposed harsh in-kind fines. Formally, the policy was meant to punish the communities underperforming in terms of state food procurement (similar to in-kind taxation) because local procurement shortfalls supposedly were a consequence of intentional, profit-seeking behavior. Using a weather-based instrument for the locality's blacklisting status, I document that blacklisting significantly reduced the present-day nightlight intensity (a proxy measure for economic development). Additional evidence points to entrepreneurship and trust as channels for this effect. My results support the notion that policies that suppress economic freedoms and disrupt social structure can have persistent negative effects on economic performance.

In Chapter 2, I use survey data to analyze the division of household work in several post-socialist countries during their democratic transition period and compare them to the advanced economies in 1994-2012. The results indicate that, while there are signs of convergence in time allocation patterns across countries, some differences persist. Female time availability, a conventional determinant of time allocation to unpaid work at home, matters significantly less in post-socialist economies, suggesting that the socialist norm for women to be responsible for the majority of household work despite being employed full-time persists in post-soviet societies throughout the transition.

Chapter 3 examines how individual labor supply responds to changes in (expected) inflation. In April-July 2022, in collaboration with ChaeWon Baek, we ran an experiment in an online labor market, Amazon Mechanical Turk, to establish a causal relationship between inflation expectations and individual labor supply in a high inflation setting. First, we use randomized information treatments to generate exogenous variation in subjective expectations about price inflation, wage inflation, and unemployment rate. Second, we investigate how these changes in expectations affect MTurk workers' reservation wages and the desired employment duration. We find that the resulting increase in wage inflation expectation significantly increases reservation wages. Higher expected price inflation rates, on the other hand, decrease reservation wages. Higher unemployment expectation increases the desired duration of employment and decreases reservation wages. These results suggest that wage-price spiral risks appear limited in the U.S. despite the high current price inflation rates.

To the unconquered.

Acknowledgments

This dissertation would not have been possible without the help of many people. First and foremost, I am deeply indebted to my dissertation advisor, Yuriy Gorodnichenko, for his guidance, inspiration, and belief in me. His encouragement to push beyond my limits and pursue my passions has been invaluable in shaping me into the researcher I am today. I would also like to thank all my advisors, Gérard Roland and Benjamin Schoefer, for invaluable advice, criticism, and encouragement. Their contributions have been essential in shaping my academic journey and research interests.

I am also grateful to many other professors within the economics department, including Barry Eichengreen, Emi Nakamura, Christina Romer, David Romer, and Jón Steinsson, for their helpful comments and advice at various stages of the Ph.D. process. A special thanks to Martha Olney for the invaluable advice about economic history research and mentoring undergraduate research assistants. Thanks to the Economic History Dissertation Group led by Martha, a semester of remote learning during the Covid-19 pandemic turned into productive work that led to the first chapter of my dissertation. Additionally, I am thankful to the participants of the UC Berkeley Macroeconomics Seminar, Student-Faculty Lunch, and Colloquium, as well as the Economic History Seminar and Lunch for their insightful comments. I also gratefully acknowledge the financial support from the All-UC Economic History Group, the Association for Comparative Economics Studies, and Tufts University.

I would like to thank all my Berkeley friends and colleagues who shared this journey with me. I am especially grateful to Dominik Jurek, Seung Joo Lee, Lukas Leucht, Joan Martinez, Jacob Weber, and David Wu. I am greatly indebted to my co-author Chae-Won Baek, who explored the intricacies of running online experiments with me to answer macroeconomic questions. I am also grateful to my friend and co-author Hanna Onyshchenko for her enthusiasm and valuable advice. I greatly benefited from discussions with Marc Dordal i Carreras and Sergii Meleshchuk. I would also like to extend my thanks to Yevheniia Yarmosh for her help and friendship throughout my Ph.D. studies. Special thanks to Berkeley undergraduate research assistants who helped me with various aspects of this dissertation: Daniel, Edgar, Emma, Haneeboy, Justin, Mary, and Ray. Additionally, I would like to thank Yevheniia Bondarenko for her help with data construction. Finally, I would like to express my gratitude to my non-Berkeley friends, including Anna, Julia, Igor, Iryna, and Taras, who have been incredibly supportive during this time.

I am immensely grateful to my parents, Yaroslava and Vitaliy, who have always been a source of support and inspiration for me. They have taught me to be curious, ambitious, courageous, and perseverant – qualities I regularly relied on while working on this dissertation and will benefit from in the future. I also want to express my heartfelt appreciation to my husband Vadym, who shared all the ups and downs of this journey with me and brightened this experience with his love and support. I am beyond grateful to have him by my side.

The last year of my Ph.D. has been especially challenging as it coincided with the Russian invasion of my home country, Ukraine. I would not have been able to complete this last mile of my Ph.D. journey without the support of my advisors, the Berkeley community, and fellow Ukrainians. I am especially grateful to Yuriy Gorodnichenko, Marianna Kudlyak, Vira Semenova, and Benjamin Schoefer who encouraged me to keep going when the work lost meaning in the first months of the war. I also want to express my deepest gratitude to the Ukrainian soldiers and volunteers who have been defending and supporting my country. Thanks to them, I was able to complete this Ph.D. journey. Now, I look forward to the moment when their mission of restoring peace and freedom in Ukraine will be accomplished.

Contents

List of Figures

List of Tables

Chapter 1

The Long-Term Consequences of Blacklisting: Evidence From the Ukrainian Famine of 1932-33

1.1 Introduction

Culture, social networks, and trust matter for economic development, but measuring their effects is notoriously difficult because these forces evolve slowly. The repressive policies of the Soviet authorities, which triggered the Great Famine of 1932-1933 in Ukraine (called the Holodomor[1]), offer an opportunity to study how ostracizing — if not eliminating — individualistically minded and entrepreneurial people, a backbone for any market economy, affected the long-run trajectory of Ukrainian communities. In contrast to other infamous cases of oppression where target groups were selected for non-economic reasons (e.g., ethnicity in the case of the Holocaust; gender and location in the case of Latin-American *mita*), the Soviet policies in the early 1930s targeted the most productive social groups. Using a novel dataset of georeferenced Soviet policies at a highly disaggregated level and weather shocks as a source of exogenous variation for the incidence of repressions, I document that, even 60 years after these disastrous Soviet policies, repressed communities are significantly lagging in development.

I focus on *blacklisting*, a Soviet policy that intended to deter market-oriented behavior through collective punishment of Ukrainian villages in 1932-33. Formally, blacklisting intended to enhance compliance with state collectivist policies (mandatory food procurement at below-market prices) by fighting the resistance of kulaks, a social group of "rural capitalists and entrepreneurs" perceived by the Soviet authorities as an "enemy class."[2]

[1] The Holodomor means "kill by starvation" in Ukrainian.

[2] Given the collective nature of the punishment, blacklisting targeted kulaks indirectly. Blacklisting was preceded by *dekulakization* (1930-31), a policy that aimed to eliminate kulaks as a social class through individual repressions, such as discriminatory taxation, property expropriation, exile, or execution.

The authorities blamed kulak-like profit-seeking behavior for procurement shortfalls and designed collective punishment to deter such behavior and antagonize peasants against kulaks. Starting in November 1932, blacklisted communities could be banned from trade and provision of crucial goods, prohibited from moving, and imposed harsh in-kind fines. The penalties were accompanied by propaganda that shamed blacklisted communities for exhibiting or tolerating kulak-like behavior. This policy both contributed to and was exacerbated by the disastrous famine in the countryside which peaked in early 1933.

This paper aims to estimate the long-term effects of blacklisting on the economic performance of Ukrainian villages during the post-1991 period, after Ukraine started transition to a market economy where entrepreneurs were no longer ostracized, and to shed light on the channels of their persistence. I hypothesize that blacklisting has negatively impacted long-term performance. Prior studies have demonstrated that demographic shocks that distort social structure unintentionally have persistent negative effects on economic performance (Acemoglu et al., 2011; Grosfeld et al., 2013). The effect of an intentional, economically motivated shock, like blacklisting, can be even larger and more persistent due to the impact of stigma on kulak attributes, such as self-reliance and individualist mindset, which could increase across generations. Such qualities are crucial for entrepreneurial development and positively contribute to economic growth (Bauernschuster et al., 2012; Doepke and Zilibotti, 2014; Gorodnichenko and Roland, 2011). In addition, blacklisting could reduce social trust, particularly towards authorities, since it was local authorities who assigned and executed collective punishments, sometimes for subjective reasons (Chen and Yang, 2019).

Estimating the long-term effect of blacklisting has proven to be difficult due to data constraints and identification challenges. To obtain data on the application of blacklisting, I create a new geospatial dataset of villages blacklisted in 1932-33 by georeferencing lists of blacklisted communities compiled by historians based on archival records. To do this, I manually match the blacklisted communities with their modern counterparts by tracing changes in the administrative division between the 1930s and 2017. This allows me to assign to blacklisted villages the spatial coordinates of their modern counterparts. I use these spatial coordinates to aggregate present-day socioeconomic outcomes on the village level. Although this approach restricts the choice of outcomes to those suitable for spatial aggregation, it helps to overcome the issue of a lack of disaggregated data on present-day economic activity.

Since blacklisting could be applied at the discretion of local authorities, my identification strategy exploits plausibly exogenous variation in the probability of blacklisting due to weather. Temporary weather aberrations are also relevant because the weather should have affected the probability that authorities would interpret a failure to meet the procurement plan as a consequence of intentional, even rebellious, behavior. Weather in Ukraine in 1931-1932 was less favorable for the harvest than in the previous years (Davies and Wheatcroft, 2016), and procurement plans did not properly account for local weather conditions. Because variations in weather can offer many potentially weak instruments, I use post-Lasso to select an optimal combination of weather instruments.[3]

Combining the novel data with the IV identification strategy, I document that blacklisting had a persistent and sizeable negative effect on economic activity in Ukraine during the post-1991 period. I use nighttime light intensity recorded by satellites as a proxy for the level of economic development (Chen and Nordhaus, 2011; Henderson et al., 2012). My IV estimates suggest that blacklisting reduced nightlight intensity in the affected villages by 1.2-1.8 log points, which corresponds to a decline in local economic activity by about 20%. The negative effect persists throughout the 20 years that I analyzed. I perform a series of robustness checks to verify that my results are unlikely to be driven by omitted factors and specification choices.

To shed light on the potential channels of persistence, I also examine the effect on other outcomes. In particular, I estimate the effects of blacklisting on the intensity of self-employment (a proxy for entrepreneurship) and voting outcomes during presidential elections (a proxy for trust in political institutions). I find a negative effect on the prevalence of entrepreneurs based on formal criteria and self-reported status. This result is in line with the notion that blacklisting reduced the number of people with entrepreneurial qualities. The persistence of the effect throughout the Soviet era could be the consequence of the inter-generational transmission of culture that maximizes one's chances of survival in case of a revival of repressions (Doepke and Zilibotti, 2014; Rozenas and Zhukov, 2019). Severe limitations of rural out-migration imposed by Soviet legislation between 1933 and 1974 may be a factor responsible for the persistence. I also find that blacklisting reduced voter turnout for the presidential elections in 2004 and 2010 which suggests that communities affected by blacklisting exhibit lower trust in political institutions. There is little support for demography and changes in ethnic composition as channels of persistence.

My paper builds on nascent literature about the consequences of the 1932-33 famine in Ukraine. Prior research has documented that population losses during the famine had long-term demographic, economic, and political effects (Naumenko, 2019; Rozenas and Zhukov, 2019).[4][5] Famine mortality can be viewed as a summary statistic for repressive

[3]Building on Belloni et al. (2012, 2014a), wide literature uses Lasso as a method to optimally reduce the dimensionality of the instruments and, thus, avoid the problem of including weak instruments. The post-Lasso procedure means that all the Lasso is used only to obtain the set of instruments that are plugged directly into the first stage of the analysis.

[4]Naumenko (2019) studies the effect of famine mortality on urban population size, industrialization, and agricultural development in Belarus, Russia, and Ukraine using panel province-level and city-level data. The main result is that despite the fact that the famine was mostly a rural phenomenon, it has a more persistent effect on urban population size. She also finds a short-term negative effect on agricultural production as well as a long-term negative effect on industrial production.

[5]Rozenas and Zhukov (2019) study the political outcomes of famine mortality which is considered a summary statistic for the incidence of mass repression. Instrumenting district-level famine mortality with local weather shocks they find that higher incidence of repression initially caused higher share of political support of the Soviet regime but this support declined over time. They attribute the changing patterns in political behavior of the affected communities to the declining power of retribution of the Soviet regime. My identification strategy builds on their approach.

collectivist policies, ethnic biases, and bad weather (Naumenko, 2021). My paper provides the first identified evidence of a very localized repressive policy that contributed to the famine. Accounting for the fact that blacklisting was assigned on the village level, I perform analysis on a much more disaggregated geographical level than done previously (district or province) which sharpens my identification. I also study the transmission of the effect through entrepreneurship, a channel that has not been previously examined in this context. Second, my paper contributes to broader literature about the long-term consequences of mass violence (e.g., Acemoglu et al., 2011; Lupu and Peisakhin, 2017; Rozenas et al., 2017; Walden and Zhukov, 2020). A unique feature of my setting is the economic motivation behind the allocation of collective punishment. Third, my paper contributes to comparative economics literature evaluating the successes and failures of the Soviet economy, particularly under Stalin's rule (Allen, 2021; Cheremukhin et al., 2017). Whereas prior literature mostly focuses on evaluating the benefits of industrialization, my paper sheds light on the economic cost of entrepreneurial and social capital destroyed in the countryside. It is also related to the literature studying transition economies (Galenson et al., 2004; Roland, 2002; Sonin, 2013). Ukraine is among a few countries that have experienced a particularly severe decline in GDP during the first decade of transition.[6] Prior literature focuses on the role of post-1991 institutions and policies to explain this phenomenon. My findings suggest that the history of collectivist repressions can be partially responsible for the downturn. Finally, my paper contributes to a broader literature about the effect of policies and institutions on long-term growth (Acemoglu et al., 2001; Dell, 2010) as well as literature about the role of culture and social structure in the economy (Acemoglu et al., 2011; Algan and Cahuc, 2014; Doepke and Zilibotti, 2014; Grosfeld et al., 2013) by providing empirical evidence about the negative effect of a policy stigmatizing and penalizing market culture on long-term economic performance.

The chapter is organized as follows. Section 1.2 summarizes the historical context for blacklisting – famine of 1932-33. Section 1.3 explains blacklisting, the targeted groups, the penalties, and application rules. The IV research design and construction of instruments is summarized in Section 1.4. Section 1.5 introduces the novel historical data about blacklisting and the main outcome variable, nightlight intensity. Section 1.6 presents the results and Section 1.7 presents the empirical evidence on the potential channels. Section 1.8 concludes.

1.2 Historical Context

This paper focuses on evaluating the effects of blacklisting in Ukraine during the famine of 1932-33. The famine was one of the greatest peacetime disasters of the 20th century: almost 11 million people perished from starvation in the Soviet Union. It was largely caused by the repressive Soviet collectivist policies in the countryside that rapidly changed a

[6] By the end of 1999, the real GDP of Ukraine dropped by 60% relative to its 1990 level (World Bank, 2002).

relatively liberal market regime of 1921-1927.[7] First, collectivization and individual repressions against the most productive peasants undermined the agricultural potential of the Ukrainian countryside. Second, throughout 1930-33, the government forcefully reallocated significant amounts of food from the countryside to the cities through the state procurement system. Overambitious state procurement plans and numerous repressive policies (such as blacklisting) used to accelerate the procurement campaign made mass starvation inevitable. Due to the man-made nature of the famine, in 2006, the parliament of Ukraine passed a law recognizing the famine of 1932-33 as a genocide against the Ukrainian people.[8]

First Five-Year Plan. The history of the famine is related to the over-ambitious first five-year plan (1928-32) that focused on rapid industrialization and collectivization of agriculture (see Figure 1.1). The plan assumed that both Soviet industry and consumption could increase simultaneously despite huge amounts of resources being transferred from consumption to investment (Hunter, 1973). The authorities not only strictly enforced the plan, but also pressured the peasants to exceed it. In the beginning of 1933, Stalin declared that the five-year plan was successfully completed in four years.[9]

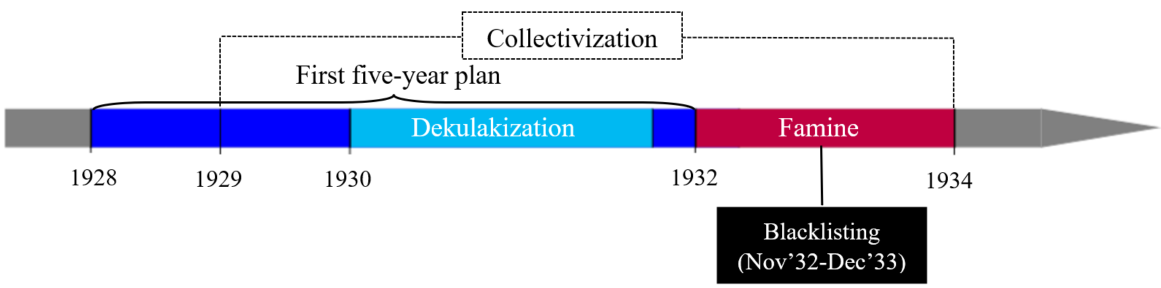


Figure 1.1: Timeline of events surrounding the famine of 1932-33

[7]After a disastrous attempt to eliminate market relations through forced food requisition and redistribution in 1919-21, which resulted in the famine of 1921-22, the New Economic Policy (NEP) was introduced. Under the NEP (1921-27), the state continued to control heavy industry, banking, transportation, and wholesale trade, whereas markets directed agriculture, retail trade, and small-scale industry. Market incentives promoted rapid agricultural growth (Gregory, 2014). In 1922-23, private trade outside cooperative and state stores constituted 75.3% of all retail trade (Davies et al., 1998). The opportunity to sell grain surpluses in the market made peasants reluctant to sell them to the state at below-market prices and endangered state procurement. To mitigate the challenges with food procurement which was essential for industrialization, in 1929 the Soviet authorities launched a collectivization campaign and effectively reintroduced the forced food requisition system of 1919.

[8]The term "genocide" was first applied to the Holodomor by the U.S. lawyer Raphael Lemkin (Lemkin, 1953). Nowadays, 17 countries (Australia, Canada, Colombia, Ecuador, Estonia, Georgia, Hungary, Latvia, Lithuania, Mexico, Paraguay, Peru, Poland, Portugal, Ukraine, USA, Vatican City) and multiple municipalities and cities recognize the Holodomor as an act of the genocide against the Ukrainian people (Holodomor Museum, 2021; Shandra, 2018).

[9]Stalin's report on the fulfillment of the plan in January 1933 is available at Marxists Internet Archive.

In 1929, the agricultural sector, a crucial part of the Soviet economy, was outside of the state's control. To enable rapid industrialization, the state needed to redistribute large amounts of food via centralized procurement. To collect enough food without openly seizing the land from peasants, the state resorted to collectivization (i.e., "socialization" of the means of production, such as land, equipment, and livestock). It would allow to substitute individual ownership with collective and make collective farms rather than individuals key actors in agriculture. The assets and products of collective farms would be collectively owned, and de-facto controlled by local authorities and party representatives.

The idea of collectivization received strong resistance among peasants, especially those with high and middle incomes, referred to as *kulaks* (or *kurkuli* in Ukrainian). As can be seen from the quote below, Joseph Stalin expected that a system of large collective farms would be as productive as kulaks' households. Disregarding the important role of kulaks' human and social capital in the rural economy, Stalin launched a full-scale repression campaign against them called dekulakization.

> Today, we have an adequate material base for us to strike at the kulaks, to break their resistance, to eliminate them as a class, and to replace their output by the output of the collective farms and state farms.
>
> *Joseph Stalin, Speech on Agrarian Policy*[10]*, Dec. 27, 1929*

Who Was a Soviet Kulak? According to Lewin (1966), there was substantial disagreement about the definition of the term "kulak" in the Soviet era. This term emerged under Tsarism, when it denoted the most efficient agricultural producers loyal to the regime. In addition to their economic role, kulaks played an important social role by being role models. Under the early Soviet regime, the term typically denoted prosperous peasants, "rural capitalists", and "peasant entrepreneurs" all of which were considered enemies of the regime.

To "strike the kulak" with discriminatory taxes and punish him with targeted repressions, clear criteria were necessary. In May 1929, Sovnarkom, the highest executive authority of the Soviet Union, defined a kulak as a person that satisfied at least one of the following criteria:[11]

1. hiring of permanent workers for agricultural work or artisan industry;
2. ownership of an industrial enterprise (such as a flour mill, dairy, fruit and vegetable drying station) if it is equipped with an engine, or a wind-mill or water wheel;
3. hiring out complex agricultural machines driven by engine either permanently or seasonally;

[10] Full text is available here: https://history.hanover.edu/courses/excerpts/111stalin.html.

[11] See Mykhaylychenko and Shatalina (1992), doc. 33.

4. hiring out of equipped premises for dwelling or business purposes permanently or seasonally;

5. presence of members in the family who are engaged in commerce or usury or who have other sources of income not derived from labor (including ministers of cults).

Those who met the criteria first faced only discriminatory taxation[12] but were later subjected to dekulakization repressions.According to different agencies' estimates, 3.9-5% of all Soviet households satisfied the definition of kulaks (Lewin, 1966). However, anti-kulak policies affected a much larger number of people. Since many poorer peasants disagreed with considering kulaks as defined by the formal criteria as an "enemy class", the term *podkulachnik* (kulak-supporter) was used to indicate anyone opposing collectivization and Soviet rule irrespective of their social status.[13] Labeling someone as kulak or podkulachnik was sufficient to justify multiple penalties and repressions toward peasants based on criteria of income or political position.

Under dekulakization of 1930-31, kulak households faced the threat of property expropriation, arrest, deportation to a labor camp, or even execution.[14] Some kulaks resisted collectivization by slaughtering livestock and destroying equipment so that it was not expropriated, or even supporting rebellions against the authorities. Others abandoned their land and villages and moved to towns to avoid the repressions. Both kinds of responses undermined the agricultural potential of the countryside (Applebaum, 2017; Davies and Wheatcroft, 2016).

Collectivization and Harvest. The first five-year plan predicted rapid collectivization and a continuous increase in agricultural output. By the end of 1932, the collectivization target was declared to be met. By October 1932, about 69% of rural households and 80% of sown land were collectivized (Appendix Figure A.1) with collectivization rates surpassing 90% in the main grain-producing areas as mandated by the state (Appendix Figure A.2). Such a high collectivization rate was reached by coercion. In the process, peasants, whether kulaks or not, slaughtered their livestock so as not to give it to the collective farms. The livestock drop started during collectivization and continued throughout the years of famine (Appendix Figure A.1).

The coercive transition to collective farming undermined agricultural productivity for several reasons. First, the redistribution of land between different categories of individual and collective users caused disruption in crop rotation. Individual farmers were endowed

[12]In 1929, tax rates for kulaks were 20 times higher than for the poorest peasants although they had only by 5 times higher income (Lewin, 1966).

[13]Viola (2000) discusses peasant letters to the newspaper *Bednota* (*Poor Peasants*) in 1924 in response to a prompt "Who is Considered a Kulak and Who [is Considered] a Laborer". In the letters, peasants disagreed that wealth alone should be a criterion for kulak status because *all* peasants strove to be wealthy. Moreover, kulaks often helped out poorer peasants (e.g., with seed and food loans). The letters agreed, however, that well-off peasants who acquired wealth in dishonest and exploitative ways should be considered kulaks.

[14]Wolowyna et al. (2016) report that 364,500 kulaks were evicted in 1930-33 in Ukraine.

with the worst land plots if endowed at all. Second, the state started providing mandatory one-size-fits-all instructions for agricultural routines that were previously managed on an individual basis. Such micromanagement made proper adjustment to local conditions impossible (Davies and Wheatcroft, 2016).[15] Third, the state did not fulfill the promise of providing collective farms with access to advanced agricultural equipment. Fourth, the collective farmers, who in the mid-1920s managed to provide food not only to feed their families but also to sell in the market, no longer had financial incentives to work effectively (Yakubova, 2011). Therefore, the agricultural output decreased in light of collectivization chaos (see Appendix Figure A.3).

In addition to collectivization, the harvest in 1931 and 1932 was affected by unfavorable weather (Davies and Wheatcroft, 2016). 1931 was the first year of unfavorable weather after a year of particularly good weather. A cold spring of 1931 led to delays in sowing and made grain development more vulnerable and warm dry winds disrupted the usual colder and wetter weather in summer. Weather conditions in 1932 with a cold spring and a hot and wet summer were even less favorable. The centralized decision-making in agriculture made counteracting the effects of adverse weather conditions challenging. Altogether, collectivization and weather resulted in harvest being lower than expected: in 1932 it was 40% below plan (Applebaum, 2017). However, the Soviet authorities refused to accept bad weather as an excuse to reduce the grain procurement plan in Ukraine.

Procurement. The state demand for food was satisfied via a centralized grain procurement system. The total amount of grain to be procured was determined based on the projected harvest which, in turn, was estimated as a product of the planned area cultivated and projected yield. The assumption of unrealistically high yields resulted in overambitious aggregate procurement plans.[16] The aggregate plan was further broken down into village-level plans, up to quotas to be contributed by each collective farm or individual peasant household. Individual peasants received higher quotas than collective farms. Kulaks received the highest rigid quotas unrelated either to land owned or the harvest.

The grain collections started around June-July as soon as harvesting started and before the official harvest was known. Mobilization brigades consisting of specially trained communist party members were sent to villages for grain collection. Members of brigades and local authorities faced administrative and legal responsibility in case of "insufficient" procurement efforts which were interpreted as "sabotage" and "collaboration with ku-

[15] Collective farmers could no longer determine how to organize work: they were required to work year-round according to the centralized agendas. The agricultural year started with the sowing of winter crops in September and ended with the harvesting and processing of spring crops in July-August (Asatkina, 1935, p. 245). Activities such as sowing, weeding, and harvesting were regulated by 5-day plans. Violation of such plans soon became subject to administrative penalties and repressions, such as blacklisting.

[16] The highest pre-revolutionary harvest in the Soviet Union was 80.1 mln. tons. The 1931 plan assumed a harvest of 98.8 mln tons and the plan for 1932, according to unpublished documents, was about 90 mln. tons (Davies and Wheatcroft, 2016).

laks." In response, many local authorities not only strictly enforced the existing plans but also pushed grain collections beyond the original plan by setting counter plans. Due to such pressure, in 1931-32, many localities exceeded the procurement plans with fewer reserves remaining to survive the upcoming famine (Appendix Figure A.4). The collections were meant to be finished by December each year but, in fact, they ended when the party approved them.[17] In some localities, the government authorized the collection of grain set aside as seed stock for the past year's procurement. Thus, fulfilling one year's plan endangered the next year's.

Famine. By 1932, most households had their grain reserves depleted. When starvation started spreading throughout the country, the government continued prioritizing rapid industrialization. Instead of adjusting unreasonably high procurement plans, authorities came up with new repressive ways to extract food from peasants – expropriation of seed stock, criminal sentences for retaining a tiny fraction of communal harvest, mandatory food searches, and requisitions. Inflexible procurement combined with repressions resulted in massive famine in the countryside (Appendix Figure A.5).

Starvation existed in multiple republics of the Soviet Union, but it was the most pronounced in Ukraine. Naumenko (2021) concludes that the primary cause of the mass starvation in Ukraine were the Soviet collectivist policies rather than the weather. Markevich et al. (2021) document that policies were designed in such a way that ethnic Ukrainians were more likely to die from starvation than people of other nationalities both within and outside Ukraine's borders. Almost 4 million people in Ukraine (13% of the country's population) perished from starvation (Rudnytskyi et al., 2015). In 1933, mortality increased 8-fold relative to the pre-famine period which exceeds the mortality during the Great Chinese famine in relative terms (Meng et al., 2015). The famine ended when the grain quotas system was substituted with a grain tax system based on the area cultivated.

1.3 Blacklisting

Blacklisting was one of the repressive policies used by the Soviet government in 1932-33. Although it was applied in several other Soviet territories, this section focuses on the Ukrainian context.

Blacklisting in Ukraine served two purposes: to accelerate grain procurement, and to stir anti-kulak sentiment. In the fall of 1932, Ukraine was lagging behind the grain procurement plan. In October 1932, the special commission headed by Vyacheslav Molotov concluded that, unlike the other Soviet republics, the sluggish fulfillment of the plan in Ukraine was not due to the lack of harvest but due to the concealing of grain and resistance of the kulaks.[18] In November 1932, Ukraine fulfilled only 60% of the procurement

[17]The 1931-32 grain collections for state procurement continued until February. As soon as it ended, collection of grain to a centralized seed stock started.

[18]According to Davies and Wheatcroft (2016), when the government initiated industrialization and increased procurement plans in 1928, peasants resisted selling grain to the state at a low price. In 1931, the

plan, justifying more radical measures for accelerating grain collections. The "kulak resistance" was to be broken by all means possible. For this purpose, blacklisting, a new repressive method combining repressions and social pressure was applied.[19]

As a method of moral stimulation, blacklisting existed in the Soviet Union before the famine. In the 1920s, individuals or organizations put on the "red board" (or list) were publicly praised for their achievements and productivity, and those put on the "black board" were shamed for their poor performance. However, before the famine, blacklisting did not threaten one's survival. In 1932, the existing Ural-Siberian method of social pressure was combined with the repressive component into a new experimental method of collective punishment.

Legislation About Blacklisting. On November 18, 1932, the Soviet government passed a secret decree No. 105 "On Measures of Strengthening of Grain Procurement" that approved applying repressions to non-compliant villages and farms. The decree blamed kulak resistance for failure to meet grain procurement plans and mandated that collective farms were put on blacklists and subject to a set of repressive measures (Stasiuk et al., 2021). The section "On methods of fighting kulak resistance" read:

> In order to overcome the kulak resistance in fulfilling the grain procurement plans, the Central Committee of the CP(b)U mandates:
>
> 1. To blacklist collective farms especially maliciously sabotaging the state grain procurement plan.
>
> With regard to blacklisted collective farms, the following measures should be implemented:
> ...

The penalties authorized by decree No. 105 included a ban on any form of trade, including withdrawal of all consumer goods from the stores, termination or expedited recovery of loans, and purges of local authorities. Multiple soviet authorities brainstormed new penalties to be introduced into blacklisting, so the list of penalties was quickly growing (Papakin, 2013).

Blacklisting Penalties. The repressive measures were meant to isolate the blacklisted communities both economically and physically. One or several penalties were usually applied: (1) ban on the supply of groceries and industrial goods; (2) forbidding trade with agricultural products; (3) acceleration of all other plans (e.g, meat or seed stock);

---

market price of rye was 61 ruble 35 kopecks while the state price was only 5 rubles 50 kopecks. The disparity for wheat was even larger. In 1926-27 (under the NEP), price differences an order of magnitude smaller made peasants reluctant to sell grain to the state. Preventing such behavior was one of the motivations behind blacklisting.

[19] Most sources on the topic of the Holodomor mention blacklisting among other repressive tools (Applebaum, 2017; Davies and Wheatcroft, 2016; Snyder, 2010). However, very few focus on blacklisting per se. My main sources of information about blacklisting are two publications of Ukrainian historians, Papakin (2013) and Stasiuk et al. (2021).

(4) in-kind fines of other produce, e.g. meat or potatoes, on top of regular collections requirements; (5) restricting access to grain and seed loans; (6) expropriating collective farms property and then making its members fulfill elevated procurement quotas as individual farmers; (7) prohibiting migration; (8) imprisoning or executing collective farms leadership and local authorities.

The ban on the provision of industrial goods, ban on migration, and purges of local authorities were the most common forms of repression (Stasiuk et al., 2021). Some of these penalties were later applied to non-blacklisted communities as well.[20] The distinguishing feature of blacklisting was the combination of multiple repressive measures with state propaganda and social pressure through indiscriminate collective punishment. Local authorities often had discretion in the intensity of repressions.[21]

Targeted Groups of Repressions and Propaganda. Blacklisting was a collective repressions policy. Entire communities were put on blacklists and all their members were affected by repressions. Such nondiscriminatory application of repressions was justified by speculation that unsatisfactory performance that triggered blacklisting decisions was driven by resistance and sabotage of kulaks, and every community member should share collective responsibility for giving in to kulak pressure. According to Papakin (2013), after passing a decree on blacklisting, authorities strategically chose large and rebellious villages with strong kulaks' establishments for exemplary blacklisting in December 1932. Therefore, among the community members affected by repressions, those possessing kulak-like attributes and expressing dissent to the authorities were to be blamed for blacklisting by other community members. Unlike the dekulakization of 1930-31 that targeted kulaks directly, blacklisting was meant to fight kulaks indirectly, through social pressure.

Evidence of Blacklisting. Many legal documents, reports to the government, and newspapers about blacklisting have survived in archives for two reasons. First, the state was monitoring the application of blacklisting. The government required local authorities to report statistics about blacklisting: whether repressions had started, what repressions were applied, and whether they helped to accelerate collections. The report "On implementation of decrees on acceleration of grain procurement and applying repressive measures, including blacklisting" from December 2, 1932, enlists numerous districts, villages, collective farms, and individuals blacklisted since November 18. Second, to strengthen the effect on the population, the government mandated that blacklisting was extensively covered in the press.[22] Figure 1.2 provides an example of a newspaper publication about blacklisted communities (villages and collective farms) and the penalties applied to them.

[20] For example, the registration system introduced in January 1933 universally banned unauthorized rural-urban migration (Kessler, 2001).

[21] E.g., some communities in the district could be subject to a complete set of repressions permitted by the law, whereas others could be subject only to partial repressions (Papakin, 2013).

[22] For example, by the end of November 1932, 10 out of 38 newspapers in the Chernihiv region included information about blacklisting (Stasiuk et al., 2021). After the communist party leadership criticized editors of the local press for limited coverage in December 1932, some journalists joined village brigades and covered news from inside blacklisted communities.

ЧОРНА ДОШКА

1. СЕЛО ПІСКИ.
2. КОЛГОСП „НОВЕ ЖИТТЯ", ЛОБРІЄВСЬКОЇ СІЛЬРАДИ.
3. КОЛГОСП „СХІД". К-МИКОЛАЇВСЬКОЇ СІЛЬРАДИ.
4. КОЛГОСП „МОДЕЛЬНИЙ ЦЕХ", Н-СЕРГІЄВСЬКОЇ СІЛЬРАДИ.
5. КОЛГОСП „14 РОКОВИН ЖОВТНЯ", БАРМАШІВСЬКОЇ СІЛЬРАДИ.
6. КОЛГОСП „ЧЕРВОНА УКРАЇНА", ПРИВІЛЬНЯНСЬКОЇ СІЛЬРАДИ.

В КОЛГОСПАХ ЧОРНОЇ ДОШКИ

1. Припинено кооперативну, державну торгівлю. Ввесь крам з крамниць, в тому числі сирники, гас вивезено до тих колгоспів, що сумлінно виконують хлібозаготівлі.

2. Накладено арешти на поточні рахунки цих колгоспів у філії Держбанку.

3. Припинено кредитування цих колгоспів. Стягується всю заборгованість достроково з колгоспів, колгоспників й тих господарств, які виключені з колгоспу за останній час.

4. За саботаж та злісне невиконання державних завдань хлібозаготівель, фінплану колгосп „М-Цех", Н-Сергієвської с-ради. оштрафовано м'ясом розміром норми 15 місячної сплати, а саме: 56,82 центн., стягнуто 55,70 центн.

Колгосп „14 річчя Жовтня" оштрафовано в 64,98 цен., „Червона Україна" в 120.16 цен., „Схід" в 71,47 цен. Ввесь штраф по цих колгоспах стягується поряд з стягненням заборгованости.

5. Заборонено колгоспникам, цих колгоспів робити перемол зерна на млинах.

6. Хліб, що одержали колгоспники в рахунок натуральних авансів зверх 10 відс. до обмолоту, стягується як не законно одержаний.

Figure 1.2: Newspaper publication with a list of blacklisted communities and penalties applied to them

Notes: From Newspaper “Under the Flag of Lenin”, issue No. 150 dd. January 1, 1933. The black frame at the top of the page titled “The Black Board” includes the list of one blacklisted village and five blacklisted collective farms in different village councils of the Odesa region. The section at the the bottom of the page titled “In the collective farms on the black board” reports penalties applied to the blacklisted communities. Specifically:

1. Cooperative, state trade has been terminated. All the goods from the shops, including matches and kerosene, were reallocated to the collective farms that conscientiously fulfill grain procurement.
2. Seizures have been imposed on the current accounts of these collective farms at the State Bank branch.
3. Lending to these collective farms has been stopped. All debts are being collected ahead of time from collective farms, current collective farm workers and previous workers who have been recently excluded from the collective farm.
4. For sabotage and malicious non-fulfillment of state tasks of grain procurement and financial plan:
   - The collective farm “Mechanic Workshop” had been imposed meat fines in the amount of 15 months’ payment, namely 56.82 quintals of which 55.70 quintals have been already collected.
   - The collective farm “14th October Anniversary” had been imposed meat fines in the amount of 64.9 quintals, "Red Ukraine" – 120.16 quintals, “East” – 71.47 quintals. The entire amount of fines for these collective farms is to be collected in addition to the collection of the existing liabilities.
5. Collective farmers of these collective farms are prohibited from grinding grain in mills.
6. Bread received by collective farmers as in-kind advances that exceed 10% of threshing is being withdrawn as illegally obtained.

Possible Consequences of Blacklisting. Blacklisting could have demographic, socioeconomic, and political consequences. There is no statistical evidence about the mortality inflicted by blacklisting (net of other factors, such as other repressive policies, ethnic bias, and weather). However, the consensus in the historical literature is that it has dramatically increased mortality.[23] The elevated mortality created a possibility for diluting the local culture by resettling residents from other Soviet republics to the areas deserted by the famine. Altogether, mortality and resettlement contributed to the eradication of the individualistic component of the pre-famine culture. The cultural shift could be further exacerbated across generations. However, if some cultural feature was disproportionately prevalent in blacklisted communities, this pattern could be preserved over time although to a lesser extent, again through intergenerational transmission of human capital and culture.[24] From a political perspective, blacklisting, just like other indiscriminate acts of violence, could both induce obedience to the state and incite opposition depending on the state's capacity (Rozenas and Zhukov, 2019).

1.4 Research Design

This section discusses my identification strategy to estimate the causal effect of blacklisting on long-term outcomes.

1.4.1 Motivation for IV

I estimate the long-term effect of blacklisting on long-term economic outcomes on the village council level using the following regression equation:

$$Y_{v,post91} = \beta_0 + \beta_1 \times BL_{v,1933} + \mathbf{X}'_{v,d}\theta_y + \varepsilon_v \tag{1.1}$$

where $Y_{v,post91}$ stands for the present-day outcome of village council v, $BL_{v,1933}$ is the indicator variable for the blacklisting status of the village council's territory in November 1932-December 1933, and $\mathbf{X}'_{v,d}$ is a set of the village council-level (v) and district-level (d) control variables.

The OLS estimate of β_1 would be unbiased if the error term ε_v is uncorrelated with blacklisting status. Given the discussion in Section 1.3, this assumption is unlikely to hold due to non-random selection into blacklisting with selection rules, although for-

[23] Papakin (2013), Snyder (2010), and Wolowyna et al. (2016) compare a decision to blacklist a particular community with a death sentence to the local population. This conclusion is made based on testimonies of eyewitnesses and reports of local authorities.

[24] Even though blacklisting increased mortality, usually at least some community members survived. According to eyewitnesses, mortality was the highest among the vulnerable groups of the population (children and elderly). The working-age population, especially those actively working in the collective farms had higher chances of survival (Papakin, 2013).

mally based on the fulfillment of the procurement plan, loosely defined.[25] Therefore, equation (1.1) is likely to suffer from the omitted variable bias, either positive or negative depending on the actual selection rule into blacklisting and how it is related to unobserved variables in the error term. If the communities with the worst fundamentals were more likely to be blacklisted (e.g., because they were the least efficient), the OLS estimate of the effect of blacklisting would be biased downward. If villages with good fundamentals were more likely to be blacklisted (e.g., because they were more entrepreneurial and had more to lose from collectivization), the OLS estimate of the effect of blacklisting would be biased upward.[26]

Having a valid instrument for the locality's blacklisting status would help to overcome the omitted variable bias. It would allow estimating the causal effect of blacklisting via 2SLS with equation (1.1) in the second stage and equation (1.2) in the first stage.

$$BL_{v,1933} = \gamma_0 + \gamma_1 Instr_{v,1933} + \mathbf{X}'_{v,d}\theta_B + \nu_v \tag{1.2}$$

In Subsection 1.4.2, I argue that a local weather shock could be a valid instrument for the locality's blacklisting status. The bases for constructing the weather shock are local weather aberrations specified as deviations of air temperature and precipitation in each month and location from the usual weather conditions in the years preceding blacklisting (defined as a median over a five-year period). In this case, the IV estimate of β_1 will capture the local average treatment effect of blacklisting on the long-term outcomes for localities that got blacklisted due to local weather aberrations captured by the weather shock. The identifying assumption is that, after controlling for other characteristics, the weather shocks affect the long-term outcomes only through blacklisting. The remainder of this section explains how I construct my preferred weather instrument.

1.4.2 Weather Shocks as an Instrument for Blacklisting

A valid instrument should satisfy two main requirements: relevance and exogeneity.

[25] On the one hand, there is evidence that fulfilling the procurement plan by 70% allowed some villages to get off a blacklist. On the other hand, local authorities often made blacklisting decisions at their discretion violating the requirement of approving them with regional authorities. Such violations could make blacklisting both more and less political.

[26] Suppose that an unobserved fundamental is positively correlated with the economic potential of a community in a market economy. Such a fundamental could be the prevalence of individualism which has been proven to positively affect innovation and economic growth. It is an omitted variable from the perspective of my regression. In this case, the long regression is $Y_{v,post91} = \beta_0 + \beta_1 \times BL_{v,1933} + \gamma EconPot_v + \mathbf{X}'_{v,d}\theta_y + \varepsilon_v$ with $\gamma > 0$. The direction of correlation between the economic potential and blacklisting, as summarized by equation $EconPot_v = \kappa + \delta BL_{v,1933} + \nu_v$ is not obvious ex-ante. However, it determines the OLS estimate bias: $b_{OLS} - \beta = \gamma\delta$. The selection into blacklisting of communities with good fundamentals ($\delta > 0$) results in an upward bias of the OLS estimate of the effect of blacklisting.

Relevance. Local weather shocks are relevant predictors of blacklisting through their effect on fulfillment of the grain procurement plans.

$$\textit{Bad weather shocks} \rightarrow \textit{Pr(Unintentional procurement shortfall)} \rightarrow Pr(Blacklisting)$$

The idea behind the collective repressions was to punish blacklisted communities for *intentionally* lagging behind the procurement plans by hiding grain from the collection brigades (kulak-like behavior) or work negligence that reduced the harvest (sabotage). Although state authorities approved food aid distribution in some localities to mitigate the famine, they did not formally recognize the unfavorable weather as an excuse for reducing the aggregate procurement plan in Ukraine (Davies and Wheatcroft, 2016). However, local authorities directly responsible for the application of repression had more reliable information about weather and its impact on harvest. Since weather was a factor outside of peasants' control, experiencing an adverse weather shock could be a mitigating factor for local authorities in allocating blacklisting decisions.

The fact that weather affects harvests is well known. For example, very hot summer weather reduces harvest (Dell et al., 2014, 2012). As discussed in Section 1.2, historians argue that bad weather reduced the harvest in Ukraine at the onset of the famine (Davies and Wheatcroft, 2016). Procurement plans were set without properly accounting for local weather conditions, so local weather shocks should affect the probability of procurement shortfalls through their effect on the harvest. Presumably, local authorities took various factors into account when applying blacklisting. Obvious forms of resistance (e.g. protests, slaughtering livestock, exits from collective farms) should have increased the probability of blacklisting whereas factors outside peasants' control that, nonetheless, led to procurement shortfalls should reduce the probability of the punishment. In this case, the decision rule for blacklisting of village v by higher-level authorities d could be viewed as a binary rule of the following form:

$$Blacklist_{v,d} = \mathbb{1}\{\text{PlanFulfill} < \underline{x}_{v,d} \;\&\; \text{Pr(Intent} \mid \textbf{weather}, \text{protests, etc.)} > \bar{z}_{v,d}\} \quad (1.3)$$

where local authorities choose the cutoffs for the deviation plan and interpret the severity of other forms of resistance at their discretion. Adverse weather shocks should be a mitigating factor in such a decision rule. Although I cannot test the relationship between weather shocks, harvest, and resistance to procurement directly, I can exploit the relationship between weather shocks and grain procurement which is available on the district level from HURI Famine Web Map data.

Exogeneity. Exogeneity requires that local weather aberrations affect the long-term outcomes only through blacklisting. Is this the case? Although it is possible that year-to-year weather changes affect economic growth (Dell et al., 2012), my choice of instruments as deviations of weather in a specific month and location from their respective median in the preceding years is plausibly exogenous to long-term outcomes. Such weather aberrations, however, could affect contemporaneous outcomes, in particular, famine severity, which has some persistent long-run effects per se (Naumenko, 2019; Rozenas and Zhukov, 2019). I cannot distinguish the famine deaths due to weather versus repressions in the data.[27] However, Naumenko (2021) concludes that weather explains no more than 8% of famine deaths whereas collectivist policies explain over 50% of the deaths. Naumenko (2021) also argues that in several other years similar weather as in 1931-32 (but without repression) did not result in famine. Therefore, the potential bias due to the direct effect of weather aberrations on famine mortality is likely quantitatively small relative to the effect via repressions. Because *transitory* local weather aberrations are unlikely to be correlated with other important events of the 20th century (e.g. WWII, Chornobyl catastrophe), this IV ensures that mass casualties and destruction resulting from these events are unlikely to confound my estimates. However, adding controls for exposure to these events may improve the precision of the results.

Threats to Identification. If some omitted factor is related to both blacklisting and instruments, my estimates could be biased. For example, instruments could be correlated with agricultural productivity if places that experience more volatile weather are less productive. I address this concern by controlling for metrics of weather variability in robustness exercises. I also control for climate zone in all specifications and examine weather patterns in the 1930s in more detail in Appendix A.2.

Two additional factors that could distort the effect of blacklisting are migration and discriminatory economic policies. Voluntary migration from the countryside was limited between 1933 and 1974 due to the mandatory registration system imposed in urban areas. The remaining out-migration was subject to state control (e.g., peasants sent to labor camps or deported due to ethnic cleansing, peasants hired to work in urban construction projects).[28] After the famine, the population of decimated villages was sometimes resettled with peasants from Belarus and Russia but the magnitude of this flow was rather small (Rozovyk, 2020; Rudnytskyi et al., 2015).[29] Overall, migration following blacklisting should attenuate my estimates due to diffusion of transmission channels. The selective migration to the cities in 1930 to avoid dekulakization could bias my results, but it was partially offset by the registration system introduced in January 1933 which meant to purge cities of kulaks (Kessler, 2001).

Discrimination of blacklisted communities though economic policies (e.g. infrastructure investment) decades after blacklisting would be costly. To ensure the stable supply of

[27] Appendix Figure A.6 provides a schematic map for the long-term effect of blacklisting and famine mortality.

[28] Detailed data about the rural out-migration is not available. To recover the regional variation in the Holodomor losses, Wolowyna et al. (2016) account for 3,085,800 rural-to-urban internal migrants in Ukraine in 1927-1938. The 11-year migrant flow is smaller than the estimated 3,942,500 excess famine deaths during the Holodomor of 1932-33. The internal rural-to-rural migration had a much smaller scope – less than 30,000 families were resettled in 1934-35.

[29] According to Rudnytskyi et al. (2015), 137,800 persons were resettled but over half of them left within the first two years.

food from the countryside, it would be necessary to integrate all rural communities into the infrastructure network and provide them with necessary equipment. Given that by the mid-1930s the Soviet regime declared a successful transition to collective agriculture, continuing discriminating against blacklisted communities would hinder food procurement and, thus, industrialization. Although I am not aware of such policies, this concern can be mitigated by adding controls that would capture such investment. Generally, my approach to addressing these and related concerns is to include controls that proxy for these potentially confounding factors (e.g., controls for climate zones account for the composition of crops which could be driving certain investment decisions). To further dispel such concerns, I will use a battery of robustness checks to explore the sensitivity of my estimates to other controls although finding some controls will require future work (e.g., village-level capital investment during the Soviet times).

Measurement Errors in Blacklisting Status. The list of blacklisted communities from Papakin (2010) was constructed based on high-level archive sources and may omit information on the blacklisting status of some units. If blacklisted communities are missing in documents available in central archives at random, mislabeling blacklisted communities would lead to attenuation bias, reduce the precision of my estimates and make it harder to detect the effect.

Interpretation. My instrument is meant to capture exogenous variation in probability of repression to evaluate the effect on present-day outcomes. In this case, the coefficient of interest β_1 captures the long-term effect of blacklisting on compliers — communities that were blacklisted due to local weather aberrations in 1931-32 and would not have been blacklisted had the weather been different.

1.4.3 Instrument Selection

Variable Specification. In line with literature that emphasizes weather anomalies, I construct local variation in weather conditions by subtracting from the average monthly temperature and precipitation in the two years preceding the famine median weather in the same locality and month during a benchmark period according to the equation below:

$$DevFromMedian^{w}_{m,y} = Value^{w}_{m,y} - Median^{w}_{m,1926-1930} \tag{1.4}$$

where $w = \{\text{temperature, precipitation}\}$, $m = \{1, 12\}$, $y = \{1931, 1932\}$. Rozenas and Zhukov (2019) use a similar set of candidate instruments to ensure removal of the systematic year-to-year or seasonal variation.[30] The periods considered are 1931-32 for weather deviations (the period right before the famine) and 1926-30 for the median. The weather

[30] Such a specification is motivated by the data available. It is similar to regressing the weather for the year, month, and day or week fixed effect as Gilchrist and Sands (2016) do to construct their instrument.

in the years 1926-30 is used as a benchmark because it characterizes the conditions in which the over-ambitious five-year plan was set.[31]

To obtain the information about the weather in the 1930s on the village level, I use the interpolated gridded dataset constructed by Matsuura and Willmott (2014).[32] It has annual information about the average monthly temperature and total monthly precipitation starting from 1900 and is commonly used in the literature (Dell et al., 2012; Naumenko, 2021; Rozenas and Zhukov, 2019). I aggregate the data at the village council level (the same level as blacklisting) by weighting it by area. The resolution of the weather data is 0.5×0.5 degrees longitude-latitude which corresponds to about 1995 sq. km. The village council area is 64.7 km on average and district area in 1933 is 2693.75 sq. km.[33] It means that one grid cell of weather data covers an area roughly equal to the area of one district or multiple village councils. Both village council-level and district-level weather variables have similar first and second moments, but distributions are smoother on the village council level due to area weighting.

Dimensionality Reduction. Equation (1.4) produces a set of candidate instruments that consists of 48 monthly deviations for temperature and precipitation in each month of 1931-32.[34] Because in practice weather in some periods a plays more important role for the harvest than in others, the set of months in which authorities considered bad weather as a mitigating factor for blacklisting is likely much smaller than 48. Therefore, I use a Lasso method to select weather-based instruments optimally.[35] This allows me to avoid the IV bias arising from the inclusion of many weak instruments (Stock et al., 2002) and increase the statistical power. Following Gilchrist and Sands (2016), Rozenas and Zhukov (2019) and others, I use Lasso only to select a subset of instruments that best predict the target variable. In other words, I plug the Lasso instruments directly in the first stage and discard the coefficients produced by Lasso. This approach is called post-Lasso. According to Belloni et al. (2014b,a), it performs well relative to alternative methods used to address the problem with many weak instruments as long as the sparsity assumption is satisfied.[36]

[31] Davies and Wheatcroft (2016) conclude that weather during the benchmark period, 1925-30, was favorable with the exception of 1927 which resulted in the famine of 1928-29. The famine was particularly severe in the main grain-producing area of Ukraine (Hrynevych, 2013).

[32] The main issues with the gridded data are the reliance on extrapolation from a limited number of weather stations, possibly leading to the attenuation bias, and spatial correlation between the variables and outcomes. Although interpolation may produce unreliable results for rugged territories (Dell, 2010), it should not be a problem for Soviet Ukraine with its relatively flat terrain. The issue of spatial correlation can be mitigated with the inclusion of controls for latitude and longitude.

[33] District area varies from 470.8 to 6647.0 sq. km.

[34] Appendix A.2 provides additional information about the candidate instruments and how they differ from the usual weather conditions.

[35] In robustness checks, I also apply several approaches (LIML and JIVE) to mitigate concerns about the bias arising from weak instruments.

[36] According to Belloni et al. (2012), when the instrument set is sparse, i.e. it can be well approximated by a small subset of the instrument, the least absolute shrinkage and selection operator (Lasso) is an effective way to select an optimal subset of instruments.

Existing literature applies Lasso directly to the endogenous variable (e.g., Angrist and Frandsen, 2022; Belloni et al., 2012; Gilchrist and Sands, 2016; Rozenas and Zhukov, 2019). Such an approach poses a concern of over-fitting the first stage. For this purpose, in my preferred specification, I apply Lasso not to my endogenous variable per se, but to the deviation from the procurement plan – an observable that should drive blacklisting according to the legislation (see Section 1.3). Formally, I apply Lasso to select instruments that best predict residualized deviation from the grain procurement plan in 1932-33 at the district level, ($\widehat{\eta_d}$), according to equation (1.6) below. The controls in equation (1.5) are meant to capture the effect of factors other than weather that should affect the district's ability to meet the procurement plan.

$$\begin{aligned} DevPlanPcnt_d =& \alpha_0 + \alpha_1 ProcurPlan_d + \alpha_2 ClimZone_d + \alpha_3 \downarrow horses_{r,1928-1932} + \\ & + \alpha_4 RailPort_d + \eta_d \end{aligned} \tag{1.5}$$

$$\widehat{\eta_d} = \sum_{m=1}^{48} \delta_m \cdot DevFromMedian_{d,m} + error \tag{1.6}$$

I choose a preferred instruments set by comparing the F-statistics for the excluded instruments in the first stage. My preferred set of instruments is:

- Deviation of air temperature from the median in May 1931, March 1932, and June 1932.
- Deviation of precipitation from the median in April 1931, January 1932, February 1932, and December 1932.

This set was obtained with a cross-validation Lasso applied to the residualized deviation from the plan as specified in equation (1.6) and produces first-stage F-statistic of 26 with just 7 instruments.[37] These weather variables only partially capture the periods indicated by historians as problematic. This is plausible because I want the instrument to pick local weather shocks rather than all-Ukrainian shocks.[38] It may seem surprising that winter matters for the deviation from the procurement plan. In fact, a large portion of wheat harvest in Ukraine was obtained from winter crops which are sown in the fall and germinate during winter. Therefore, winter weather could plausibly affect the harvest.

1.4.4 Correction for Spatial Correlation

A common concern in persistence studies is that observed empirical results may be driven not by a true structural relationship but by a spurious spatial relationship. This may affect both the estimates of the coefficients and standard errors. To avoid the omitted variable

[37] See other candidate instruments in Table A.1. My preferred set of instruments is in column 4.

[38] Many conclusions of Davies and Wheatcroft (2016) are based on weather information in the Kyiv region.

bias due to the correlation of the explanatory variable of interest with some unobserved spatial shocks, I control for the latitude and longitude of the administrative unit's centroid in my preferred regression specifications (Lind, 2019). In addition to the omitted variable bias, there is a concern that spatial correlation may invalidate the statistical tests due to incorrect estimation of standard errors. Two approaches that are typically used in the literature are a) clustering standard errors on the administrative unit level, and b) computing Conley standard errors (Conley, 1999).[39] I use standard errors developed by Colella et al. (2019) to correct for spatial correlation.[40] In my preferred specifications, I report Conley standard errors corrected for spatial correlation within 50 km village council's centroid.[41] An alternative approach would be to cluster standard errors on the district level which typically produces slightly lower standard errors than the Conley correction.

1.5 Data

The unit of observation for my analysis is a village council. The village council designates the territory of one or several villages governed by a common administration. The key historical variable is the blacklisting status of the village council's territory in 1932-33, which I created by georeferencing lists of blacklisted communities created by historians. Additional historical data are discussed in Section 1.2 and Table 1.1 and historical weather data are discussed in Section 1.4.3. Present-day (post-1991) outcomes include variables characterizing economic activity and other outcomes that are suitable for spatial aggregation. Administrative units in Ukraine are rather small. The population for only 183 of the 6,983 administrative units that were part of the Soviet Union in 1933 exceeded 20,000 in 2001.[42] Therefore, in the analysis, I restrict attention to the remaining 6,762 administrative units to capture the fact that famine and blacklisting was a mostly rural phenomenon.

1.5.1 Blacklisting

After the Soviet Union collapsed and archives opened, historians compiled lists of locations that were subject to blacklisting in the 1930s. In this project, I georeferenced lists of blacklisted communities created by historian Heorhii Papakin. Papakin (2010) pro-

[39] See Online Appendix for Dell et al. (2014) for an overview.

[40] The Stata package *acreg* corrects standard errors for an arbitrary form of correlation. Arbitrary means that each observation's error term may depend on another observation's error term to a certain degree. In the case of Conley (1999) correction, the dependence is summarized by the distance between observations. In the case of standard errors clustering, Colella et al. (2019) relax the assumption of correlation only within non-overlapping clusters imposed by Cameron et al. (2011) which is used in Stata's *ivreg* command.

[41] The circle with this radius covers the grid cell for which weather data is reported as well as the average district area in 1933.

[42] The village-level population count according to the 2001 Census is publicly available at Census Data Bank but it is not georeferenced. The administrator of the website *Datatowel.in.ua* kindly shared georeferenced data with me.

vides information about the location of blacklisted communities (collective farm, village, or village council) according to the administrative division of Ukraine in the 1930s but without distinguishing the penalties applied to the communities. It was created based on legal documents and press publications available in the state and regional archives of Ukraine. It covers the entire territory of Ukraine, but it is not exhaustive. Some evidence of blacklisting may have been destroyed, or it may be available only in local archives.

Since their publication, Papakin's lists of blacklisted communities attracted a lot of attention. In 2013, Harvard Ukrainian Research Institute published schematic maps based on these lists indicating the presence of blacklisted communities somewhere in the district.[43] Because the blacklisting was applied at the village level, district-level variation is too coarse to identify the effect of blacklisting.[44] To study the consequences of blacklisting, it is critical to tie the lists to specific villages.

Assignment of the spatial coordinates to the units in Papakin's lists is a challenging task for several reasons. First, to my knowledge, there is no village-level map of Ukraine from the early 1930s. Second, the administrative division of Ukraine changed many times between 1932 and 1991. Many villages were renamed, dissolved, moved to different districts, etc.[45]. Third, there is no database of collective farms on the village level in the early 1930s. Furthermore, collective farms were being created and dissolved very often at the time.[46] Because some collective farm names are very common, it is impossible to infer the village in which they were located from the collective farm name.[47]

To address these challenges, I georeferenced Papakin's lists by manually tracing the names of each blacklisted community over time until I found their modern counterparts. I used information from handbooks about administrative division, regional Memory Books of the Holodomor victims, websites of local administrations, and other sources to create a historical record for each village. Eventually, I merged the list of blacklisted communities with polygons for the village councils in Ukraine in 2017 obtained from geoBoundaries project (Runfola et al., 2020). The resulting map of blacklisted localities is provided in Figure 1.3.[48]

According to the 2017 administrative division, there were 10,357 local councils representing 29,726 villages, towns, and cities in Ukraine. Of them, the territory of 6,993 local councils with 21,463 villages, towns, and cities was a part of the Soviet Union and, thus, hit by the famine of 1932-33.[49] Since I will restrict the analysis to smaller administrative

[43]The maps are available at https://gis.huri.harvard.edu/blacklisted-localities.

[44]In 1933, each district included on average 21 village councils, each of which consisted of several villages (Asatkina, 1935).

[45]I was unable to match villages that disappeared shortly after the famine to present-day counterparts. In the Mykolaiv region alone, 128 villages, both blacklisted and non-blacklisted, disappeared before 1946.

[46]Blacklisting could be a reason for the collective farm dissolution. In this case, their property would be transferred to the state farms' ownership.

[47]E.g. "October's", "Lenin's", "Stalin's", "Communist" collective farm.

[48]This new database should be a valuable resource for researchers interested in studying the Holodomor and will be made available online.

[49]Local council consists of one or several administrative units. It corresponds to the third level of ad-

units which eliminates most of the cities and towns, I refer to the unit of observation as the village council rather than local council.[50] I focus on village councils rather than villages in order to reduce the probability of erroneously attributing the blacklisted unit to villages that were not blacklisted. Because village boundaries may have changed multiple times since the 1930s, the probability of a mistake is smaller for the village councils because they include several villages.

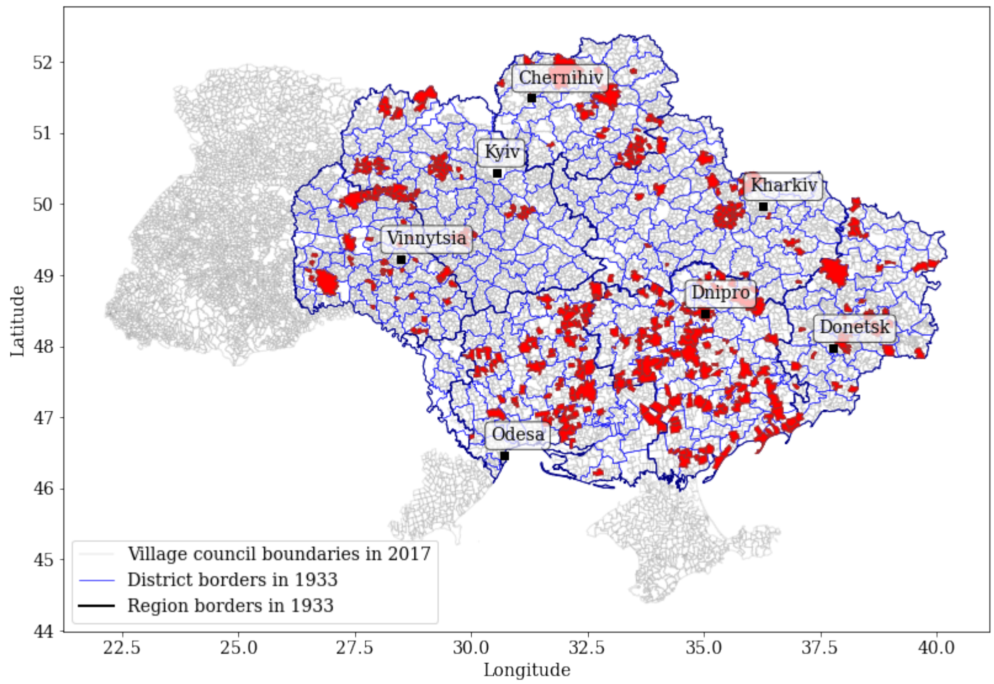


Figure 1.3: Map of communities blacklisted between November 18, 1932 and December 31, 1933 according to the georeferenced lists.

Notes: The map is based on the list from Papakin (2010). Blacklisted communities are denoted with red, and not blacklisted ones are denoted with white.

Papakin's raw lists include 1,285 blacklisted units (rows) which I further process by removing duplicates and units for which a modern counterpart was not found. I was able to match 1044 unique blacklisted units with 599 modern village councils. Figure 1.4 summarizes the number of cases of blacklisting in the raw dataset by date of blacklisting. Although Papakin (2010) provides information about some units blacklisted in 1931 and 1934, I restrict attention to those blacklisted between November 18, 1932, and December 31, 1933 – the time when blacklisting with repressive components was actively used to

ministrative division of Ukraine: 1. oblast (region or province), 2. raion (district or county), 3. village or city council, 4. village or city.

[50]Cities and towns were much less affected by the famine because the Soviet government wanted to protect the urban population engaged in industrial production.

accelerate grain procurement and before the procurement system became more flexible. After decree No. 105 was passed and until the end of 1933, the territory of at least 416 (6%) modern village councils in Ukraine was blacklisted.

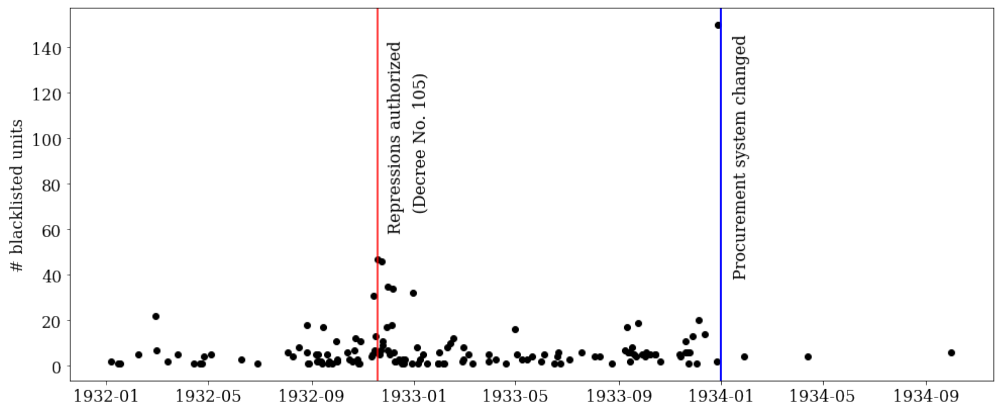


Figure 1.4: The number of blacklisting cases by date according to the georeferenced lists from Papakin (2010)

Notes: The actual number of blacklisting cases may be different because information about blacklisting of some units may be mentioned in multiple sources at different moments in time.

Correlation With Other Historical Variables. Additional historical variables, such as grain procurement plans and famine mortality, are available at the district level mostly from the Harvard Ukrainian Research Institute (HURI) Famine Web Map and replication package for Naumenko (2021). As can be seen in Table 1.1, in 1927, districts with blacklisted units were on average smaller, less densely populated, and had fewer Ukrainians, and more Russians and Germans. There is no significant difference in the share of Jews. Districts with blacklisted units also had more equipment per capita and about the same amount of livestock per capita in 1925 and access to a railroad or port from the district center. They had higher grain procurement plans (in tons) and were lagging behind the plans more severely. Districts with blacklisted units also had smaller famine losses, probably due to smaller population size. The intensity of famine losses was also smaller in districts with more blacklisted units but the difference was not statistically significant.

The absence of a positive correlation between blacklisting and district-level famine mortality seems to contradict the claim of Papakin (2013) that blacklisting meant effectively a death sentence, but selection into blacklisting can obscure the causal effect. Enforcement of blacklisting required allocation of the scarce resources (food for collection brigades' members). Therefore, local authorities had an incentive to apply blacklisting to communities without apparent signs of mass starvation to increase probability of extracting large amounts of food there. In addition, according to historians, local authorities were advised to apply blacklisting as a "punishment of the last resort" to communities that did not respond to other forms of repression.

Table 1.1: District-level characteristics depending on the presence of blacklisted villages on their territory

Variable	Not blacklisted units		Difference with blacklisted units (δ)	
	Mean	SD	Mean	S.e.
Population (1927)	81134.9	[52185.0]	-8253.809	(4349.7)
Rural population density (1930)	25.70	[9.73]	-0.319	(1.94)
Share of Ukrainians (1927)	87.10	[14.9]	-1.975	(1.80)
Share of Russians (1927)	5.00	[8.84]	1.208	(1.05)
Share of Jews (1927)	1.70	[2.04]	0.091	(0.26)
Share of Germans (1927)	2.05	[6.30]	0.706	(0.71)
Livestock per capita (1925)	0.47	[0.12]	0.005	(0.02)
Equipment per capita (1925)	0.071	[0.036]	0.014	(0.01)**
Direct access to railroad or port (1933)	0.365	[0.481]	0.003	(0.06)
Grain procurement plan, tons (1932-33)	11973.2	[11104.0]	55357	(2164.4)**
% fulfillment of plan (1932-33)	-0.09	[0.14]	-0.071	(0.02)***
Rural famine losses, abs (1933-34)	9492.7	[7006.4]	-1007.2	(922.7)
Rural famine losses, per 1000 (1933-34)	149.5	[96.93]	-2.23	(13.14)

Notes: The table reports results from regression $Y_d^v = \alpha + \delta \times BL_{v,d} + \varepsilon_v$ where Y_d^v stands for the district-level outcome. The constant corresponds to the mean value for districts without blacklisted villages weighted by the number of villages, and the coefficient corresponds to the weighted average difference in outcomes of districts with and without blacklisted communities. Standard errors are corrected for spatial correlation within radius of 50 km: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Data sources: HURI Famine Web Map and replication package for Naumenko (2021).

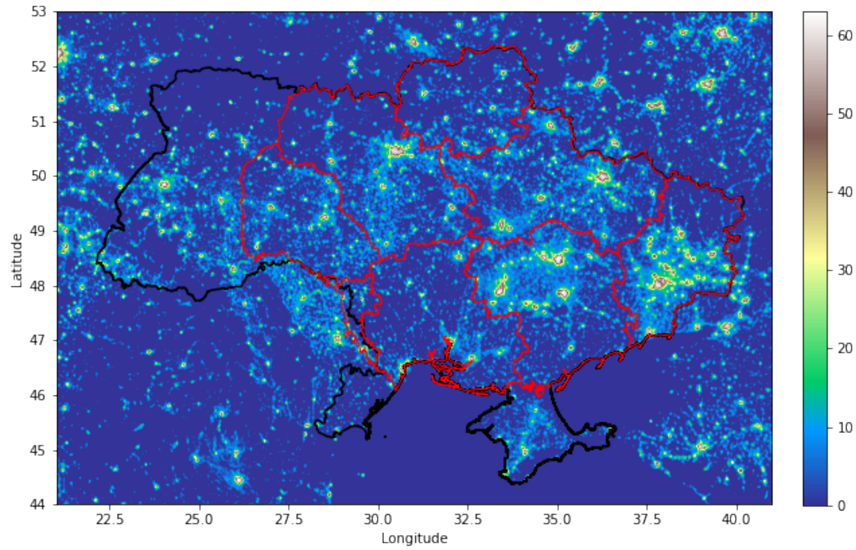
1.5.2 Present-Day Outcomes

Nightlight Intensity. To measure economic activity on the local level, I use the gridded data of nighttime light intensity obtained from satellite images.[51] Specifically, I focus on cloud-free composites for average visible stable night-time lights obtained from satellite images in 1992-2013.[52] Panel (a) of Figure 1.5 plots the nighttime light intensity for 1992 and 2013. The brightness of nighttime lights takes values 1-63 per pixel with background noise set to 0. Each pixel corresponds to a 30-arc-second grid cell and covers an area of about one square kilometer.

[51] This proxy is motivated by the assumption that usage of nighttime lights per person increases in income because most consumption and production activities in the evening require lights. While geographically disaggregated income data may be not available to researchers, nightlights are (Henderson et al., 2012).

[52] Data is constructed by National Oceanic and Atmospheric Administration's National Geophysical Data Center and US Air Force Weather Agency. A number of data-cleaning procedures were applied to the raw nightlight data in order to remove all the sources of light except for electricity (sunlight, moonlight, clouds, aurora). The data for 1992-2013 is available here: https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html. For every year, the products are 30 arc second grids, spanning -180 to 180 degrees longitude and -65 to 75 degrees latitude.

(a) Ukraine



(b) Andriivska village council in Chernihiv oblast

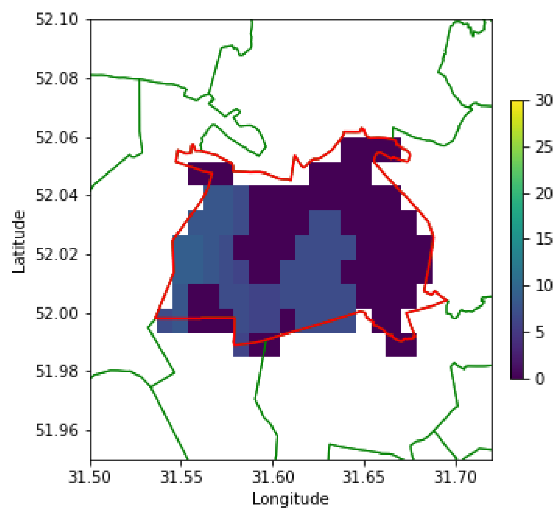


Figure 1.5: Nighttime light intensity in Ukraine in 1992

Notes: The figure plots cloud-free composites for average visible stable nighttime lights constructed by NOAA's National Geophysical Data Center and US Air Force Weather Agency. Each pixel corresponds to an area of about one square kilometer. The brightness of the average visible stable lights takes values 1-63 per pixel with background noise set to 0. Black line denotes the present-day border of Ukraine. Red lines in Panel (a) denote borders of regions in 1933. Red and green lines in Panel (b) denote borders of village councils in 2017.

To aggregate the data on the administrative unit level, I compute the total amount of nightlight on the polygon level, as a sum of the brightness of each pixel of the administrative unit's territory. Then, I normalize the total nightlight intensity by the number of pixels to account for the fact that larger units may be brighter merely due to their size. If for a given year, the data are available for multiple satellites, I take the mean of the average nightlight intensity per well-lit pixel across satellites to obtain a unique value for each year. Panel (b) of Figure 1.5 visualizes the nightlight intensity in one village council of Ukraine used to calculate the average nightlight intensity per well-lit pixel.

There is a substantial number of pixels in different years with missing data on nightlight intensity. The most likely reason is the absence of cloud-free satellite images. If I dropped observations with missing data in at least one year, it would reduce my sample size by up to 1000 observations. To maximize the number of observations, I impute missing values as the average of nightlight intensity of the immediate neighbors of the same type (village, town, or city).[53]

Table 1.2 reports summary statistics for the imputed nightlight intensity. The average value of the brightness of nightlight per well-lit pixel in a village council is between 5.74 and 8.80, which corresponds to 1.69-2.12 log points. Administrative units that were blacklisted appear to be on average only slightly dimmer than those that were not blacklisted.

Table 1.2: Summary statistics for nightlight intensity by administrative unit's blacklisting status

Variable	Not blacklisted		Blacklisted - not blacklisted		N
	Mean	SD	Mean	S.e.	
NL per pixel, 1992	8.80	[3.84]	-0.23	(0.18)	6093
NL per pixel, 2001	5.74	[2.76]	-0.08	(0.13)	6092
NL per pixel, 2012	7.70	[4.50]	-0.59	(0.14)*	6093
Log (NL per pixel), 1992	2.12	[0.31]	-0.01	(0.02)	6093
Log (NL per pixel), 2001	1.69	[0.30]	0.00	(0.02)	6092
Log (NL per pixel), 2012	1.96	[0.35]	-0.04	(0.02)	6093

Notes: The table reports results from regression $Y_v = \alpha + \delta \times BL_v + error_v$ where Y_d^v stands for village-level outcome. The sample consists of units with fewer than 20,000 inhabitants in 2001. Standard errors are corrected for spatial correlation within radius of 50 km: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I consider the logarithm of nightlight intensity per well-lit pixel as the main regression outcome. To understand how the effect on nightlights maps to economic activity, I estimate the elasticity of real gross regional product with respect to the nightlight intensity per well-lit pixel in 2001-2013 (see Appendix A.3 for details). For economic interpretation, I will use a range of elasticity estimates: 0.13 as a lower bound and 0.31 as an upper

[53] An alternative imputation approach would be to replace missing values in a given year with the sample mean for rural areas and urban areas respectively. My results are not sensitive to a change in imputation procedure.

bound.[54] These estimates suggest that an increase in nightlight intensity per pixel by 1 log point corresponds to an increase in gross regional product by 0.13-0.31 log points.

Prior literature points out that nightlight works better as a proxy of economic activity on the national rather than local level and that it tends to underestimate the economic activity in urban areas and overestimate it in rural ones (Doll et al., 2000; Ghosh et al., 2010; Ishizawa et al., 2019; Mellander et al., 2015). It suggests that the elasticity of output to nightlights can be different for urban and rural areas. Because I do not have data on gross regional product with a breakdown by urban/rural status, I will refer to the aggregate elasticity when interpreting the results. Given that I focus on cross-sectional variation, restrict the analysis to locations with less than 20 000 residents (mostly rural areas) and analyze the nightlight intensity per pixel, these concerns should not significantly undermine my results. Moreover, interpretation in terms of units of nightlight intensity is robust to measurement errors in terms of elasticity.

1.6 Results

This section discusses the results about the effect of blacklisting on economic activity using the data and research design described earlier. Section 1.6.1 summarizes the first stage and provides reduced form evidence about the effect of weather on nightlights. Section 1.6.2 reports the main IV results and verifies their robustness to inclusion of alternative controls and specifications.

1.6.1 The first stage, reduced form, and balance of observables

This subsection sets the background for the main IV result by providing evidence about the first stage, reduced form, and covariates balance.

First Stage. The first stage and reduced form for my preferred set of instruments are summarized in Table 1.3. Columns 1-3 indicate a strong first stage with F-statistic exceeding 26. The coefficients on the excluded instruments are stable across specifications, as suggested by p-values for equality of coefficients across columns. The weather deviations that produce the strongest first stage are temperature aberrations in March and June 1932 – the months of bad weather shocks according to Davies and Wheatcroft (2016). The positive coefficient on March 1932 temperature suggests that localities that faced cold spring weather (bad for the harvest) were less likely to be blacklisted. The negative coefficient on June 1932 temperature also suggests that localities that faced hot summer weather (bad for the harvest) were less likely to be blacklisted. These results indicate that communities less affected by the bad weather shocks were less likely to be blacklisted.

[54] 0.13 is Ukraine-specific estimate of elasticity from my analysis after accounting for all fixed effects (see column 5 of Table A.8). 0.31 is estimate of elasticity for low and middle income countries from Henderson et al. (2012).

One explanation for the negative correlation is that local authorities could use very bad weather shocks as an excuse for failure to meet procurement plans without facing retaliation from the higher-level authorities. Another explanation is the tradeoff the authorities faced when deciding whom to blacklist and the motivation to target localities where they were more likely to extract food (see Section 1.3).

Table 1.3: First stage and reduced form for the Lasso-chosen instruments

	First stage (Depvar: BL)			Reduced form (Depvar: $log(NL/pxl'92-95)$)		
	No controls (1)	Lon-lat controls (2)	All controls (3)	No controls (4)	Lon-lat controls (5)	All controls (6)
Deviation of temperature from median						
May'31	0.012	0.009	0.007	-0.055*	0.073^{b}	0.101*
	(0.022)	(0.029)	(0.027)	(0.032)	(0.063)	(0.056)
Mar'32	0.092***	0.083***	0.056***	-0.056*	0.008^{b}	-0.018
	(0.021)	(0.023)	(0.019)	(0.033)	(0.032)	(0.035)
Jun'32	-0.041***	-0.061***	-0.048*	0.086***	0.182^{***a}	0.187***
	(0.009)	(0.018)	(0.025)	(0.022)	(0.056)	(0.055)
Deviation of precipitation from median						
Apr'31	-0.003	-0.003	-0.003	0.017	0.010	0.011
	(0.006)	(0.006)	(0.006)	(0.014)	(0.013)	(0.011)
Jan'32	0.001	0.002	-0.003	-0.007	-0.015	-0.017
	(0.011)	(0.011)	(0.010)	(0.022)	(0.022)	(0.019)
Feb'32	-0.041***	-0.043***	-0.034***	0.039	0.067***	0.067***
	(0.014)	(0.014)	(0.012)	(0.025)	(0.025)	(0.025)
Dec'32	-0.008	-0.008	-0.002	0.057**	0.030	0.038**
	(0.008)	(0.008)	(0.009)	(0.024)	(0.021)	(0.019)
N	6094	6094	6094	6094	6094	6094
R-sq	0.03	0.03	0.04	0.12	0.14	0.19
F-stat	45.38	42.28	26.4	48.94	13.12	20.55
P-val		0.98	1.00		0.18	0.96
Controls	base	spatial	econ+sp	base	spatial	econ+sp

Notes: The table reports estimates of coefficients k for the excluded instruments based on equation $LHS_v = \gamma_0 + \sum_{w,M,Y} k_{w,Y,M} Instr^{w}_{v,M,Y} + \mathbf{X}'_v\theta_B + \nu_v$. Instruments are the deviation of monthly temperature (precipitation) in year Y and month M in a given village from median temperature (precipitation) in the same village over the preceding 5-year period. The sample consists of units with fewer than 20,000 residents in 2001. Standard errors are corrected for spatial correlation within radius of 50 km: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Lon-lat controls: longitude and latitude. F-stat is statistic from a test that coefficients on excluded instruments are jointly equal to 0. P-val is the p-value of the test that coefficients in a given column are the same as in the previous column. P-values for statistically significant differences in coefficients across specifications: ^{a} $p < 0.10$, ^{b} $p < 0.05$, ^{c} $p < 0.01$. All controls: district-level rural population density, livestock and equipment per capita, and access to railroad or port, indicators for the climate zones, centroid longitude and latitude.

Reduced Form. Columns 4-6 of Table 1.3 report the reduced form for the nightlight intensity per well-lit pixel in 1992-95. The F-statistics for the excluded instruments exceeding 20 suggests that these instruments have good power at predicting the outcome. Importantly, the coefficients have the opposite sign in the reduced form regressions and in the first stage: weather shocks that increased the probability of blacklisting also reduce the nightlight intensity in the long run. This result suggests that blacklisting negatively affected economic activity.

Binscatters. Panels (a) and (b) in Figure 1.6 show that blacklisting probability predicted using the preferred set of weather instruments is positively correlated with actual blacklisting status and, as expected, negatively correlated with the deviation from the grain procurement plan: villages in the districts that were more severely lagging behind the plan were more likely to be blacklisted. Panels (c) and (d) in Figure 1.6 preview the main IV result according to which blacklisting predicted using the preferred set of weather instruments significantly reduces nightlight intensity per well-lit pixel.

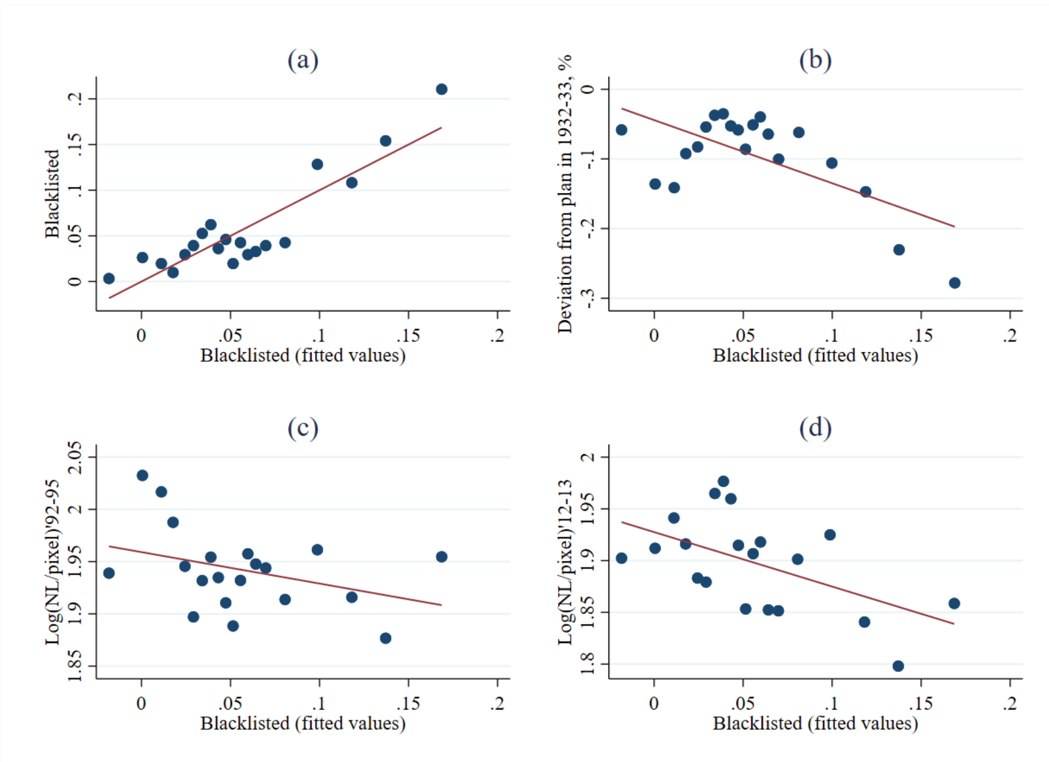


Figure 1.6: Binscatters for predicted blacklisting status and (a) actual blacklisting, (b) deviation from the grain procurement plan in 1932-33, and (c)-(d) nightlight intensity per well-lit pixel in 1992-95 and 2012-13

Notes: The figure illustrates relationship between the probability of blacklisting and other variables (without other controls). The deviation from district-level procurement plan is calculated as $\frac{Actual-Plan}{Plan}$. Sample: Administrative units with fewer than 20,000 residents are included.

Balance Table. The use of IV makes the sample less balanced in terms of some covariates (e.g., population size and density, the share of Ukrainians, livestock and equipment per capita, see Appendix Table A.2 for details). I address these imbalances by including rural population density, livestock and equipment per capita, as well as access to a railroad or port as controls in my baseline specification.[55] In robustness checks, I will verify whether the results are robust to controlling for other characteristics (population size, the share of Ukrainians, fulfillment of procurement plan in 1930-31).

1.6.2 Second Stage

To analyze the effect of blacklisting on economic activity, I estimate equations (1.1) and (1.2) for the logarithm of nightlight intensity per well-lit pixel. Table 1.4 reports the estimates together with Conley standard errors adjusted for spatial correlation within 50 km from the village council's centroid.

Table 1.4: Effect of blacklisting on nightlight intensity

	Log(Nightlight intensity per well-lit pixel)					
	1992-1995	1996-1999	2000-2003	2004-2007	2008-2011	2012-2013
Panel A: IV						
Blacklisted	-1.707	-1.401	-1.527	-1.343	-1.795	-1.164
	(0.627)***	(0.581)**	(0.613)**	(0.618)**	(0.661)***	(0.654)*
Panel B: OLS						
Blacklisted	-0.005	-0.006	-0.008	-0.023	-0.026	-0.026
	(0.022)	(0.020)	(0.020)	(0.019)	(0.020)	(0.021)
N	6094	6094	6093	6094	6092	6093
Mean outcome	1.94	1.74	1.63	1.42	1.76	1.9
S.d. outcome	0.32	0.30	0.29	0.30	0.32	0.33

Notes: The table reports estimate of β_1 from equations (1.1) and (1.2) for administrative units with fewer than 20,000 residents in 2001. First-stage F-statistic for excluded instruments is 26.4. Conley standard errors (50 km) in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls: Rural population density, livestock and equipment per capita, and access to railroad or port, climate zone, centroid longitude and latitude.

Main Result. The results in Panel A of Table 1.4 suggest that blacklisting has a persistent negative effect on nightlight intensity during the transition period. It reduced the nightlight intensity of communities blacklisted due to weather shocks by 1.2-1.8 log points. The large estimates suggest that in absence of blacklisting, economic activity in blacklisted territories could be much higher. Specifically, given a plausible elasticity of output with respect to nightlight intensity 0.13 (see Section 1.5.2 and Appendix A.3), the IV estimates for years 1992-95 in Table 1.4 imply that economic activity declined by about 22% ($\approx 0.13 \times 1.707 \times 100\%$) due to blacklisting. Elasticity of 0.31 implies decline in economic

[55] The contribution of these variables to my results is summarized in Appendix Table A.3.

activity by 52%. Repeating the same calculation for the IV estimate of -1.16 in 2012-13 implies that blacklisting reduced economic activity by 15-36% during this time. Taking into account possible measurement errors that attenuate my IV estimates, the true effect is likely larger than the conservative lower-bound estimates.

Figure 1.7 reports the moving-average IV estimates over a four-year period. All estimates are negative and large, consistent with the results in Table 1.4. The size of the estimates varies between -1.3 and -1.8 log points. I observe two distinct patterns in Figure 1.7: i) a gradual decline in coefficients before 2000; ii) a noticeable increase in 2008. The gradual decline in the effect of blacklisting during the transition to market economy is not surprising and is in line with the hypothesis about the role of anti-market propaganda during blacklisting. The increase in the effect of blacklisting in 2008, at the onset of the global financial crisis, may be a consequence of the fact that blacklisting persists through channels that are more powerful in recessions. Obviously, this interpretation is tentative but, if true, it also suggests that blacklisting can affect how sensitive communities are to economic crises.

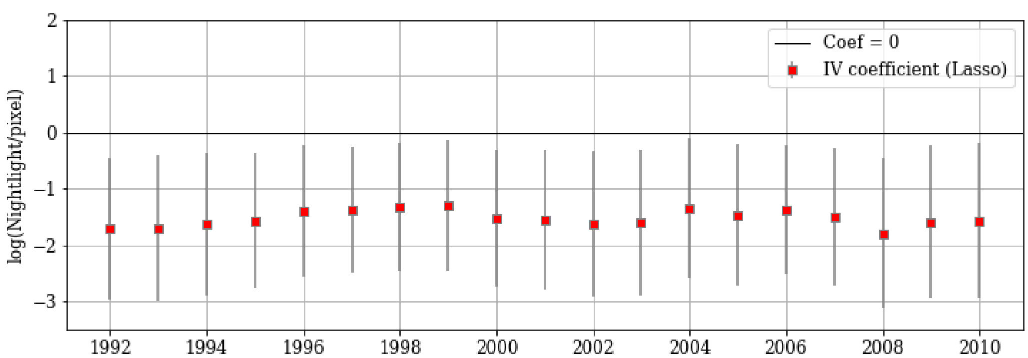


Figure 1.7: Effect of blacklisting on nightlight intensity (4-year moving average)

Notes: The figure plots IV estimate of β_1 from regression equation (1.1) where village council's blacklisting status is instrumented by weather shocks selected optimally by Lasso. The outcome is the logarithm of four-year moving average nightlight intensity per well-lit pixel. The F-statistic for the excluded instruments exceeds 26.

OLS results reported in Panel B of 1.4 are negative but very small and statistically insignificant. A comparison of the two panels indicates that OLS results are biased upward. The reason can be omitted variable bias due to missing a variable positively related to blacklisting and future economic performance, such as entrepreneurship or an individualist mindset (see discussion in Section 1.4.1). The economic costs of blacklisting must have accumulated over time.

Robustness Checks. Earlier, I discussed several factors potentially correlated with blacklisting that could pose a threat to my identification strategy. In Table 1.5, I test whether the inclusion of these factors in the model substantially affects my results. Note that the coefficient in column 1 of Table 1.5 is smaller than the one in Table 1.4. The reason for this is that some variables (collectivization rate in 1930 and fulfillment of procurement plan in 1930-31) are missing for some districts and I keep the sample consistent across all specifications.

Even though Table A.2 suggests that IV sample is not balanced in terms of pre-famine population, controlling for population size does not change the IV estimate of blacklisting. Anti-Ukrainian bias is a possible reason for the famine intensity (Markevich et al., 2021; Naumenko, 2021). Since controlling for the share of Ukrainians in the district only slightly reduced the estimate of the blacklisting coefficient, the observed effect of blacklisting is unlikely to be explained by the anti-Ukrainian bias.This result, however, should not be interpreted as no anti-Ukrainian nature of this policy. Ukrainians constituted a vast majority of rural population and with relatively little variation in the data my approach could be unable to detect any effect of anti-Ukrainian policies.

Table 1.5: Robustness of the effect on nightlight intensity to inclusion of controls

	(1) Base	(2) Pop	(3) %Ukr	(4) Collect	(5) %Procur	(6) Pooled	(7) mo_SD	(8) lon-lat2
Blacklisted (1933)	-1.415	-1.431	-1.356	-1.303	-1.180	-1.057	-1.446	-2.007
	(0.548)***	(0.542)***	(0.500)***	(0.579)**	(0.465)**	(0.446)**	(0.656)**	(0.789)**
Rural population in (1930)		-0.000				-0.000		
		(0.000)				(0.000)		
Share of Ukrainians (1927)			-0.001			-0.001		
			(0.001)			(0.001)		
Collectivization (1930)				-0.154		-0.131		
				(0.069)**		(0.060)**		
Fulfill. grain pr. plan (1930-31)					0.004	0.004		
					(0.002)**	(0.002)**		
N	5714	5714	5714	5714	5714	5714	5714	5714
F-stat (1st stage)	19.69	18.96	21.29	21.07	20.54	21.92	11.34	16.20

Notes: The table reports estimate of β_1 from equations (1.1) and (1.2) as well as coefficients on selected controls for administrative units with fewer than 20,000 residents in 2001. Conley standard errors (50 km) are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The regression specification in column 1 is the same in column 1 of Table 1.4. Columns 2-5 add one additional control at a time, and column 6 pools them together. Column 7 controls for standard deviations of monthly temperature and precipitation in months used in my instruments. Column 8 adds second order polynomial of latitude and longitude.

Collectivization in 1930 was reached mostly through voluntary collectivization so it can be viewed as a proxy for individualist culture. The variable is negatively and significantly associated with nightlight intensity consistency with evidence in the literature that individualism positively affects growth. The IV estimate of the effect of blacklisting slightly declines with the inclusion of this variable in the model. Fulfillment of the procurement plan in 1930-31 can be viewed as a proxy of laziness or resistance to state procurement. The higher rate of the fulfillment of the procurement plan is positively correlated with long-term economic performance, and controlling for it somewhat reduces the negative effect of blacklisting. Pooling previously considered controls in the same specification does not affect the main result much. In addition, the results in column 6 of Table 1.5 point to a significant correlation of post-1991 local economic activity with the collectivization rate in 1930 and the percentage fulfillment of the grain procurement plan in 1930-31.

Column 7 of Table 1.5 controls for standard deviations of each temperature and precipitation variable in the instrument set to mitigate the concern that higher deviations of weather from the median may be a consequence of permanently higher weather variability. This has little effect on the main IV result. Controlling for higher-order longitude-latitude polynomials in column 8 does not change the main effect qualitatively. It increases in magnitude and remains negative and statistically significant.

Overall, the inclusion of additional controls does not change the estimate of β_1 qualitatively: it remains negative, large, and statistically significant pointing to the persistent effect of blacklisting on long-term economic activity.

Instrument Choice. Angrist and Frandsen (2022) raise concerns about using Lasso for instrument selection when the sparsity assumption is violated. They suggest that other methods robust to weak instruments, such as limited information maximum likelihood (LIML) and jackknife instrumental variables estimator (JIVE) may be a better strategy if the candidate instrument set is substantially smaller than the sample size. Table 1.6 reports IV estimation results for nightlight intensity in 1992-1995 using these methods. Both methods indicate that blacklisting instrumented with 48 weather shocks has a large negative effect on nightlight intensity. I interpret it as an indicator that Lasso is an appropriate method of instrument selection in my case and helps to avoid the weak IV problem.

Table 1.6: Robustness of the effect of blacklisting on nightlight intensity to alternative IV methods

Log(NL per pixel'92-95)	(1) IV	(2) JIVE	(3) LIML
Blacklisted	-1.543***	-2.387***	-5.891
	(0.300)	(0.616)	(3.513)
N	6092	6092	6092
F-stat (1st stage)	6.91	7.43	6.91

Notes: The table reports estimate of β_1 from equations (1.1) and (1.2) for administrative units with fewer than 20,000 residents in 2001. Robust standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.7 Channels

The results in Section 1.6 indicate that blacklisting has a negative effect on economic activity 60 years after application. There are a number of channels through which the effect of blacklisting could persist, starting from demography (e.g. population loss, changes in ethnic composition, health deterioration) and human capital (death or imprisonment of the most skilled population) and ending with culture and trust. My hypotheses were that the effect of blacklisting persists over time because it stirred negative sentiment toward kulak attributes and, possibly, trust. To shed light on the plausibility of these mechanisms of persistence, I examine the effect of blacklisting on two sets of outcomes – entrepreneurial activity and election results. I also discuss whether other channels could rationalize the observed effect on economic activity.

1.7.1 Entrepreneurship

If the repressions associated with blacklisting created negative sentiment against kulak attributes, this could manifest itself in the efforts to set up a business after the dissolution of the Soviet Union. To test whether this is the case, I use the administrative data from the Unified State Register of Legal Entities and Individual Entrepreneurs (USR), which covers the universe of legal entities and individual entrepreneurs in Ukraine. The register contains information about individual entrepreneurs registered since 1991, but their consistent registration started only in 2003, when the registry was created.[56] The raw data contains 5,721,955 records from 1991 to 2021. To analyze entrepreneurship at the village council level, I georeferenced the registration address of each entrepreneur and mapped it to the modern village councils.

The register does not provide information about income or the number of employees but it provides information about the activities of entrepreneurs. The most popular activity kinds for individual entrepreneurs in Ukraine are retail trade (both in stores, stalls, and markets), hair salons, taxi services, car repair, market research, and computer programming. After excluding from the data the activity types that are likely to represent paid employment, I end up with 82% of the initial sample of individual entrepreneurs (4,711,722 records).[57] In addition, I omit observations where the registration date is missing or recorded incorrectly (less than 1% of the total). I was able to georeference about

[56]The records of registration of individual entrepreneurs prior to this date were likely transferred from other registries.

[57]Individual entrepreneurs in IT make up a substantial share of entrepreneurs (almost 14%). I exclude computer programming from analysis because in the IT industry regular employees often register as individual entrepreneurs to simplify the logistics of working with international clients and optimize taxes and hence it is not entrepreneurship in a traditional sense (a risky, profit-oriented activity). For similar reasons, I exclude market and public opinion research and business consulting services (they make up almost 4% of all individual entrepreneurs).

90% of all the records for entrepreneurs registered since 2004.[58] For the regression analysis, I computed the total number of individual entrepreneurs and the number of days they were active (i.e. the number of days till termination or till December 2021).

The number of individual entrepreneurs registered between 2004 and 2021 varies between 0 and 319,320 in all administrative units, and between 0 and 23,281 in units with fewer than 20,000 residents. In small administrative units, there are on average 76 individual entrepreneurs per 1000 population. I take a logarithm of this variable to make the distribution less skewed. Summary statistics for the entrepreneurial activity is available in Table 1.7.

Table 1.7: Summary statistics for entrepreneurship and voting outcomes by blacklisting status

Variable	Not blacklisted		Blacklisted - not blacklisted		N
	Mean	SD	Mean	S.e.	
Number of ind. entrepreneurs	146.6	[562.36]	54.85	(27.00)**	5770
Ind. entrepreneurs per 1000	76.03	[712.69]	-20.67	(12.8)	5770
Log(Ind. entrepreneurs per 1000)	3.72	[0.70]	0.04	(0.04)	5770
Days active	1484.50	[502.3]	15.04	(36.24)	5770
Voter turnout, 2004	79.70	[7.80]	-1.88	(0.73)**	5141
Voter turnout, 2010	73.72	[8.04]	-2.00	(0.65)***	5141
Pro-Western votes, 2004	62.40	[31.00]	-7.36	(3.94)*	5141
Pro-Western votes, 2010	52.81	[25.21]	-6.09	(-3.15)*	5141

Notes: The table reports results from regression $Y_v = \alpha + \delta \times BL_v + error_v$ where Y_d^v stands for village-level outcome. The sample consists of units with fewer than 20,000 residents in 2001. Conley standard errors (50 km) are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Columns 1 and 2 of Table 1.8 report the IV estimates for the logarithm of the number of registered individual entrepreneurs per 1,000 residents in Ukraine between 2004 and 2021 and the duration of their activity according to equations (1.1) and (1.2). Column 1 indicates that blacklisting has reduced the logarithm of the number of individual entrepreneurs per 1,000 residents. The point estimate of 2.04 corresponds to a decrease in the number of entrepreneurs by almost 3 standard deviations or 55% relative to the mean. The estimate is statistically significant at 5% but accompanied by relatively large standard errors. A conservative estimate based on the lower bound of the two-standard-deviation confidence interval implies that blacklisting reduced the incidence of entrepreneurship by over 3%.

The second column indicates that in the blacklisted localities individual entrepreneurs stay in business longer. One interpretation of this result is that in an environment where it

[58] To verify the reliability of the data, I compared the number of business entities from the georeferenced data with the official statistics published by the State Statistics Committee (available on the regional level for 2014-2019). The correspondence between the results from the two data sources is almost one-to-one.

is very uncommon to be an entrepreneur, e.g. because the population is very risk-averse, only the most confident individuals with foolproof business ideas set up a business. Alternatively, in environments with very few competitors those who happen to set up business for whatever reason stay active longer than in environments where competition drives the least effective entrepreneurs out of business.

Registered entrepreneurs are based on a formal legal definition while in practice people can engage in entrepreneurial activities without formal registration, which is especially likely in rural areas. To obtain an alternative measure of entrepreneurship, I use the data from the 10% sample of the 2001 Census on self-employment. This is the only census conducted in independent Ukraine and the information is available at the district (rather than village) level. I constructed the outcome variables by computing the share of individuals in each district reporting a certain income source as the main one – self-employment (either owning a business or working on an individual basis), farm ownership, or work on a subsidiary plot.

Table 1.8: Effect of blacklisting in entrepreneurship

	Administrative data		Census data: Main income source, %		
	log(ind. ent. per 1000)	days active	self-empl. (non-farm)	own farm	work on agr.plot
	(1)	(2)	(3)	(4)	(5)
Blacklisted	-2.040**	1493.729**	-6.753	1.231*	0.386
	(0.968)	(646.111)	(4.791)	(0.743)	(18.887)
N	5770	5770	5564	5564	5564
E[Y]	3.718	1485.4	1.931	0.361	13.539

Notes: Table reports the IV estimate of β_1 from equation (1.1). First-stage F-statistic for excluded instruments in 24.6. Dependent variables in columns 1-2 are on the village council level, and in columns 3-5 – in district level. Conley standard errors (50 km) are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls: Rural population density, livestock and equipment per capita, and access to railroad or port, climate zone, and centroid longitude and latitude.

Columns 3-5 of Table 1.8 report the estimates for the effect of blacklisting on the share of individuals with various income sources. It indicates that in districts with more blacklisted villages there is a lower non-farm self-employment rate and a higher farm ownership rate. Both point estimates are very large economically (e.g., the estimate in column 3 implies the decline in self-employment rate by 3 times due to blacklisting) but since they come with large standard errors, I cannot reject a hypothesis that blacklisting has a more moderate effect as implied by results in column 1. At the same time, blacklisting has no significant effect on deriving income mostly from work on a subsidiary plot which could be viewed as a proxy of poverty. These results imply that blacklisting reduces only the probability of self-employment but does not significantly affect the choice of other income sources.

It is worth noting that in theory, there could be reverse causality between entrepreneurship and economic activity. In this case, the decline in entrepreneurship would be a consequence of a decline in economic activity. However, prior literature documents that poor aggregate economic performance in post-socialist countries promoted self-employment due to limited alternative income sources (Fritsch et al., 2014). If entrepreneurship was a consequence of a decline in economic activity (rather than a channel) it would be very unlikely to observe the negative effect of blacklisting on entrepreneurship.

1.7.2 Trust

If repressions affected attitudes and political trust, this could be reflected in voting outcomes (Alesina and Fuchs-Schündeln, 2007). I analyze the results for presidential elections in 2004 and 2010, each with clearly defined pro-Western and pro-Russian candidates.[59] In 2004, pro-Western candidate, Viktor Yushchenko, won, in 2010 pro-Russian Viktor Yanukovych won.[60] The outcomes are voter turnout and the share of votes for the pro-Western candidate. As indicated in Table 1.7, in both years voter turnout and share of pro-Western votes are smaller in blacklisted localities.

Table 1.9: Effect of blacklisting in election results

	Turnout, %		Pro-western votes, %		Turnout, %		Pro-western votes, %	
	2004	2010	2004	2010	2004	2010	2004	2010
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Blacklisted	-5.75	-9.19	35.18	51.68	-26.73**	-23.74**	-45.21	-21.03
	(21.89)	(20.03)	(65.07)	(50.38)	(13.08)	(11.98)	(42.49)	(32.30)
N	5141	5141	5141	5141	5141	5141	5141	5141
F-stat	16.42	16.42	16.42	16.42	33.65	33.65	33.65	33.65
Climate control	Yes	Yes	Yes	Yes	No	No	No	No
E[Y]	79.59	73.60	61.96	52.44	79.59	73.60	61.96	52.44

Notes: Table reports the IV estimate of β_1 from equation (1.1). Conley standard errors in parentheses and clustered standard errors are in brackets: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls: Rural population density, livestock and equipment per capita, access to railroad or port, climate zone indicator (in columns 1-4), and centroid longitude and latitude.

The estimation results are reported in Table 1.9. They suggest that blacklisting decreased the turnout during the presidential elections, but has not significantly affected the share of votes for the pro-Western candidate both in 2004 and 2010. The point estimates

[59]Starting point for my analysis is data with georeferenced addresses of polling stations which Yuri Zhukov kindly shared with me. It allows me to aggregate the election results on the village council level.

[60]In short, Yanukovych was widely viewed as a pro-Soviet candidate. One example of a difference between Yushchenko's and Yanukovych's focus is their attitude to the Holodomor. Yushchenko viewed the Holodomor as an act of genocide against the Ukrainian people, supported research about the Holodomor, and put the reference about it on the president's website. Yanukovych, oppositely, rejected the anti-Ukrainian nature of the Holodomor and referred to it as a "tragedy of all Soviet countryside". When Yanukovych became a president, he deleted the link about the Holodomor from the president's website (Motyl, 2010).

in columns 1 and 2 imply a decrease in turnout by 6-9 percentage points, and in columns 5 and 6 a reduction by 24-27 percentage points. Table 1.9 also shows that the inclusion of climate zone indicators as controls changes the significance and direction of some results quite substantially. The climate zone indicators are meant to capture the grain-producing areas and potentially differential development paths they experienced.[61]

The absence of consistent and statistically significant results on the support of pro-western candidates contrasts with Rozenas and Zhukov (2019) finding that higher famine mortality (as a proxy for repression severity) increased the prevalence of anti-Russian votes. I interpret it as an indicator that blacklisting reduced the trust to the authorities in general, but has not affected political preferences. Such a reaction is plausible given that blacklisting decisions were often assigned and implemented by the *local* authorities. This conclusion is in line with my hypothesis that blacklisting led to the destruction of social capital and undermined political trust.

1.7.3 Other Channels and Future Directions

There are other channels that could potentially drive the persistence of blacklisting effects. This subsection discusses population size, ethnic composition, human capital and investment in infrastructure as possible channels. For some of them I cannot provide detailed evidence due to data constraints.

Rural Population Size. According to Papakin (2013), blacklisting was likely to significantly increase famine mortality (see Section 1.3). If so, the population losses could drive the persistence of the effects of blacklisting on present-day outcomes. To test the validity of this channel of persistence empirically, I apply the IV strategy discussed in Section 1.4 to the rural population count on the village council level in 1959, 1989, and 2001.[62] Due to the wide confidence intervals, the IV results do not allow me to reject either positive or negative effects of blacklisting on rural population size (see Appendix Table A.4).

Change in Ethnic Composition. Ethnically-motivated repressions and resettlement could disproportionately hit the blacklisted areas. To see whether this is the case, I analyze the effect of blacklisting on the change in district-level share of several dominant nationalities between 1927 and 2001 (see Appendix Table A.5). I do not find a significant effect of blacklisting on any nationality except for Poles. It likely captures the fact that the areas in the west that experienced resettlement of Poles after WWII were also disproportionately hit by blacklisting. There is also suggestive evidence that blacklisting contributed to the increase in the share of Ukrainians. One explanation is that it could be driven by the centralized resettlement policies, such as the resettlement of Ukrainians from Poland. Another explanation is that, consitent with Lupu and Peisakhin (2017),

[61] Results in columns 5-8 do not include any grain suitability controls. However, when I control for the share of arable land used for grain cultivation instead of climate zone indicators (not reported here), the results are very similar to those in columns 5-8.

[62] The average size of the rural population declined from 2,353 (in 1959) to 1,663 (in 1989) and 1,462 (in 2001).

people who lived in blacklisted communities between the 1930s and nowadays are more likely to self-identify as Ukrainians irrespective of their ethnic background.

Human Capital. Repressions against kulaks were a direct shock to the village-level human capital, and anti-kulak propaganda that accompanied blacklisting could dampen the rate of entrepreneurial human capital accumulation in the years that followed. If so, it could explain the decline in entrepreneurship discussed in Section 1.7.1. Blacklisting could also reduce the human capital of the affected localities through the delayed building of schools and other constraints to education access. I do not have village-level data to test this channel. However, narrative evidence about the history of blacklisted villages after the 1930s does not suggest that this is an important concern. (Such histories mention school improvements and distinguished teachers.) Changes in human capital could be also related to migration, but the power of this channel was attenuated by mobility constraints between 1933 and 1974.

Investment in Infrastructure. If Soviet authorities continued discriminating against blacklisted communities after 1933 through public investment in infrastructure (such as roads, agricultural equipment, electrification, etc.), this could rationalize the long-term economic stagnation. I do not have detailed data to rule out this explanation empirically. However, I consider it unlikely because such discrimination would be difficult for the planners and costly for the economy. Many infrastructural works (e.g. electrification) are organized by geographic principle so that neighboring units enjoy improvements at about the same time. Delaying the timing of work for one unit in the middle would be inefficient. Moreover, Soviet economy benefited from providing villages with better agricultural equipment and integrating countryside into transport networks to redistribute food from cities to villages.

1.8 Conclusions

This paper examines the long-term consequences of blacklisting, a Soviet collective repression policy applied during the famine of 1932-33 in Ukraine. Any resident could be subject to repressions inflicted by the failure of the collective farm or village community to meet the grain procurement plan, or other demands of state or local authorities. The application of indiscriminate repressions was justified by the necessity to punish behavior that presumably caused procurement shortfalls. Specifically, authorities used collective punishment to discourage market-oriented behavior associated with kulaks, relatively well-off and entrepreneurial peasants who were stigmatized by Soviet propaganda. Formally, blacklisting had two goals – accelerating the fulfillment of state procurement plans and antagonizing peasants against the kulaks.

How does blacklisting affect the present-day economic activity in Ukraine? This question is difficult to answer because the selection into blacklisting was nonrandom and the criteria could be correlated with unobservable fundamentals that affect long-term economic performance. Because local weather conditions should have entered the local authorities' decision rule on blacklisting, to overcome the identification challenge, I instrument the locality's blacklisting status with local weather shocks defined as monthly aberrations of temperature and precipitation in two years preceding the famine from usual weather conditions in the locality. The preferred instrument selected by post-Lasso produces a strong first stage and is plausibly exogenous to the long-term economic performance.

My IV estimates suggest that blacklisting has a persistent negative effect on the long-term economic performance (approximated by nightlight intensity per well-lit pixel) by 1.2-1.8 log points. It corresponds to a decline of local output by about 20% in 1992-2010. This result is likely attenuated by measurement errors in blacklisting status. The negative persistent effect is consistent with evidence in the literature about the legacy of unintended shocks to social structure (e.g., Acemoglu et al., 2011), but the magnitude is larger. This can be attributed to the fact that blacklisting, unlike other shocks examined in the literature, was designed to distort social structure. I also find negative effects of blacklisting on the incidence of entrepreneurship and voter turnout during the presidential elections. These results provide support for entrepreneurship and political trust being the primary sources of persistence. They are qualitatively similar to the consequences of persecution and mass murder of Jews – a highly skilled and most entrepreneurial part of the population in urban areas of the Soviet Union (Acemoglu et al., 2011; Grosfeld et al., 2013). This is not surprising, because in many ways kulaks played a similar economic and social role in the countryside as Jews played in urban areas.

My results support the conclusion that policies that suppress economic freedoms and disrupt social structure can have persistent negative effects on economic performance. They also shed new light on policies that emphasize the "big-push" strategy (e.g., Soviet industrialization) involving forced reallocation of resources from one group or sector to another. Specifically, although the Soviet leaders boasted rapid growth of industrial output, the massive losses in rural areas were ignored. My evidence suggests that the calculus can be profoundly affected not only by staggering mortality in the 1930s but also by the highly persistent decline in economic development of the areas affected by blacklisting and other Soviet repressive policies. My results also suggest that even more innocuous policies stigmatizing, oppressing, or exploiting highly productive groups of society (e.g. through excessive taxation) could have similar effects so that long-term losses outweigh the short-term gains.

Chapter 2

Intra-Household Time Allocation: Evidence from the Post-Socialist Countries

2.1 Introduction

Unpaid work at home is an important component of a household's life. Time use surveys document that unpaid work at home — which includes routine activities like cooking, cleaning, laundry as well as caring for family members — takes at least as much time as work for pay (Aguiar and Hurst, 2016; Kan et al., 2011). In a seminal paper, Becker (1965) attracted attention to optimal time allocation between work for pay and home production. Since then, a lot of research in economics and sociology has been dedicated to understanding determinants and consequences of time allocation within households. The consensus view is that both household-level factors (such as relative income, employment status, presence of children) and macro-level factors (such as income inequality and gender pay gaps, nature of welfare systems, social norms) matter for the intra-household decisions about time allocation.

A considerable body of empirical research provides evidence about cross-country differences in allocation of time to household work. These studies raise policy-relevant questions and examine division of household work along both extensive margin (who typically performs what tasks) as well as intensive margin (how much time household members devote to different activities). At the same time, evidence on the intra-household time allocation in post-socialist countries is scarce. This is an important gap in the literature given that the socialist countries devoted significant attention and resources to ensure that women had equal access to jobs and career growth. How this legacy affects the current allocation of housework is a key question.

State-socialist countries experienced different evolutionary trajectories relative to Western countries. Previous research documents a gradual convergence in the amount of time allocated by men and women towards household work in Western countries in the 1960s-1980s (Kan et al., 2011). It happened at the same time as gender convergence in labor force participation rates (Dvorkin and Shell, 2015). Under state socialism, male and female labor force participation rates converged much earlier due to state-provided incentives and propaganda. However, gender equality in contribution to household work, although promised by some socialist governments did not occur. Therefore, women in state socialist countries ended up working "double shifts" – in addition to being employed full-time outside the home they remained responsible for most of the domestic work (Einhorn, 1993; Saxonberg, 2014).

The democratic transition of state socialist countries that began in the late 1980s has changed both political and socio-economic environments in which households operated. First, the end of the socialist rule meant that the degree to which the state regulated the personal life trajectory decreased, and ideological pressure for women to combine working full-time outside the home as well as at home was alleviated (LaFont, 2001). Second, various market reforms were implemented which resulted in structural economic changes and unraveling pay inequality (Brainerd, 2000). Finally, the transition to a market economy has improved access to consumer durables (e.g., washing machines) which could reduce the cost of home production. Overall, the transition has affected societies in complex interrelated ways that could change the time allocation within households.

This paper examines gender patterns in time allocation to different activities in post-socialist countries and advanced economies during the post-1991 period. Such analysis generates insights about differences across countries and welfare regimes that have been under-researched previously. Given that survey data about time allocation under socialism is scarce, comparing countries early during the transition period may provide useful information for economists and policy-makers. To this end, I analyze time allocation in married couples from a set of post-socialist and advanced economies during 1994-2012 using the data from the "Family and Changing Gender Roles" module of the International Social Survey Programme (ISSP). I discuss trends in the data regarding household work allocation along both extensive and intensive margins. Then, I examine the determinants of intra-household time allocation via regression analysis. Finally, I apply the Kitagawa-Blinder-Oaxaca decomposition to better understand factors driving the similarities and differences between post-socialist countries and western economies.

I find that despite multiple differences prior to the 1990s, patterns in women's employment and time allocation were quite similar in post-socialist countries and advanced economies. The similarities in averages, however, mask some important differences persisting during the transition. The analysis suggests that post-socialist countries appear to diverge from the global trend of gender convergence. The time allocation in post-socialist countries appears to reflect the socialist-era norm that women should both work for pay and perform household work. Regression analysis indicates that an increase in married woman's relative input to market work is associated with a decrease in her relative input to household work by about the same amount across welfare regimes. At the same time, an increase in woman's relative contribution to market work is associated with a decrease of her relative input to housework by a lesser amount in post-socialist countries than in advanced economies. These differences between post-socialist and advanced economies cannot be unambiguously attributed to either observable characteristics or differences in their importance.

While many papers feature post-socialist countries in cross-country analysis of intra-household time allocation, very few focus on comparison of post-socialist countries with other policy regimes (other than comparing East and West Germany). Fuwa (2004) and Mikucka (2009) are notable exceptions. I extend their analysis through examination of a longer time horizon, inclusion of both extensive and intensive measures of time allocation, and comparison of post-socialist countries with known welfare policy regimes in a regression framework. Therefore, my paper provides more comprehensive analysis of trends and determinants of intra-household time allocation in post-socialist countries relative to western economies than prior literature does.

The chapter is organized as follows. Section 2.2 provides literature review. Section 2.3 summarizes historical, political, and socioeconomic context for analysis. Sections 2.4 and 2.5 describe the data and main trends in intra-household time allocation in 1994-2012. Section 2.6 summarizes methods for regression analysis and the Kitagawa-Blinder-Oaxaca decomposition and Section 2.7 presents the results. Section 2.8 discusses the main findings and concludes.

2.2 Literature Review

Trends and determinants of time allocation to unpaid work at home in Western countries have received a lot of attention in prior literature. The main trend is a gradual convergence in the amount of time devoted to household work by men and women which started in the 1960-1980s (Coltrane, 2000; Kan et al., 2011). Multiple factors were documented to explain the allocation of household work at individual, household, and country levels.

Theories of Time Allocation. From a sociological perspective, individual-level and household-level factors are often combined into three groups according to theories about the nature of household work division – time availability, relative resources, and gender ideology (Bianchi et al., 2000; Coltrane, 2000). Time availability theory predicts that household members working more for pay contribute less to unpaid work at home because the market work takes precedence in time allocation (e.g., Coverman, 1985). Relative resources theory considers household work division as a bargaining problem and assumes that members with more resources (summarized by income share and education) have more bargaining power, and, therefore, can avoid performing domestic work (e.g., Bittman et al., 2003; Brines, 1994). Gender ideology refers to norms, beliefs, and preferences that affect gender norms on the division of chores (Davis and Greenstein, 2009). Despite rich empirical evidence, there is no consensus on whether any of the theories is more important than the others.

Economic theories consider time allocation as an outcome of a utility maximization problem subject to time and income constraints. As a solution to this problem, time allocation to both paid and unpaid work within a household is a function of wages and other parameters (e.g., productivity parameters or penalties for deviation from social norms). A dominant economic theory relies on the principle of comparative advantage according to which women traditionally specialize in household work because they are relatively more productive in this activity and/or paid relatively less than men (Becker, 1991). While recent studies provide little support for the validity of the comparative advantage theory (e.g., Siminski and Yetsenga, 2022), there is evidence about the role of gender norms in household work allocation motivating high-earning women to perform more domestic work (e.g., Akerlof and Kranton, 2000; Bertrand et al., 2015; Bertrand, 2020).

Unlike the sociological models where the number of hours devoted to market work is treated as exogenous, a distinct feature of economic models is the assumption that time allocation to market and household work is determined simultaneously, which results in different predictions about the optimal time allocation. However, in real life, both assumptions are likely oversimplifying. With limited flexibility of market hours worked, the choice of hours worked for pay may indeed precede the choice of hours devoted to domestic labor. In this case, the time allocation to household work would be affected by the time allocation to market work in addition to pay and preference parameters. For the sake of generality, empirical specification in this paper studies the association between time allocation to household work and three groups of factors as determined by sociological theories.

Country-Level Factors. Aggregate factors that affect intra-household time allocation include the national policies related to work, family, and women's employment, labor market regulations, rewards, and culture (Baxter and Tai, 2016; Blau et al., 2020). These factors can be correlated with the country's welfare regime – a concept that summarizes the policies and culture regarding social equity, state intervention in welfare provision, and gender ideologies (Esping-Andersen, 1989). Prior literature documents that allocation of household work significantly differs across welfare regimes (Fuwa, 2004; Kan et al., 2011). In a study of time allocation to household work from the 1960s to 2000s across four policy clusters, Kan et al. (2011) document that the most dramatic decline in the relative woman's input to household work occurred in Nordic countries and countries from a liberal policy cluster (Australia, Canada, UK, USA) — from 90-95% in the 1960s to less than 75% in 2000s, and the smallest decline was observed in the Mediterranean countries. Despite gender convergence in allocation of household work across welfare regimes, gender segregation in household work persists until today. Women mostly devote time to routine tasks whereas men usually do non-routine, non-core domestic work, even in the countries with prevalent egalitarian attitudes, like Nordic ones. The paper does not feature post-socialist countries (except for Slovenia, which is considered a part of the conservative policy cluster).

Post-Socialist Countries. Although several papers feature both post-socialist and other countries, to my knowledge, none focuses on the division of household work in post-socialist countries relative to other regimes. Fuwa (2004) includes seven post-socialist countries[1] in the analysis of task sharing[2] in 22 industrialized countries in 1994. She finds that even though former socialist countries had household work almost as equally divided as social-democratic countries despite lower aggregate measures of gender equality, the factors underlying the three theories had the weakest effect on household work division in the post-socialist countries. Ukhova (2020) extends the analysis of task sharing to the 1994-2012 period and restricts analysis only to post-socialist countries. She concludes that inequality in task sharing has not significantly declined since the transition period. In a cross-regime study, Mikucka (2009) concludes that post-socialist countries stand out as a unique combination of high female workload in absolute terms with high support of traditional gender roles and egalitarian division of household work in relative terms.

In a study of the effects of state communism on household work allocation in Germany a decade after the reunification, Geist (2009) finds that after controlling for individual, family, and labor market characteristics, women in West Germany devote more time to domestic work than women in East Germany. Lippmann et al. (2020) conclude that four decades of exposure to socialism was sufficient to reverse the male breadwinner role in East Germany relative to West Germany, and this resulted in more equal allocation of household work across genders.

My paper discusses the trends in division of household work in several post-socialist countries and analyzes what factors determine the intra-household allocation in the post-socialist countries relative to advanced economies. A larger sample of countries as compared to Geist (2009) allows me to draw more generalizable conclusions, although at a cost of a larger number of underlying cross-country differences. Similar to Ukhova (2020), I analyze time allocation over time in post-socialist countries, but I also incorporate cross-regime comparisons similar to Fuwa (2004) and Mikucka (2009). Like Fuwa (2004) and Ukhova (2020), I integrate three theories about determinants of time allocation in the analysis and I analyze trends in the task-sharing index, but in regression analysis I focus on the relative contribution of women to household work along the intensive margin.

[1]Fuwa (2004) considers four welfare regimes: social democratic (Norway and Sweden), conservative (Austria, Italy, Japan, the Netherlands, West Germany), liberal (Australia, Canada, Great Britain, Ireland, New Zealand, Northern Ireland, and the United States), and former socialist (Bulgaria, Czech Republic, East Germany, Hungary, Poland, Russia, and Slovenia). The former socialist countries have the lowest gender equality scores in the sample.

[2]The task-sharing index summarizes who — a woman, a man, or both equally — is mostly responsible for doing an array of household activities typically perceived as female responsibility, such as cooking, cleaning, laundry, care for sick family members, and grocery shopping.

2.3 Historical, Political, and Socioeconomic Context

Under state socialism, gender convergence in labor force participation started much earlier than in Western countries (Brainerd, 2000; Ghodsee and Mead, 2018). However, convergence in time devoted to household work did not occur and domestic sphere remained predominantly the female's responsibility (Einhorn, 1993; Saxonberg, 2014).

Because cross-country data about female labor force participation under socialism is scarce, I refer to retrospective evidence from survey data. Figure 2.1 summarizes female labor force participation inferred from the share of respondents who said their mother worked for pay before the respondent turned 14. Given that most respondents were adults at the moment of the survey (median age in the sample is between 44 and 49 years old in each wave), their childhood occurred before 1989. In the post-socialist countries, on average 70-80% of respondents had a mother working during their childhood, and in the advanced economies, there were only 45-55% of such respondents. However, there is also substantial heterogeneity within the socialist group with the highest mothers' employment rates in the post-Soviet countries (Russia and Latvia) and East Germany, and the lowest rates in Slovenia, Hungary, and Poland.

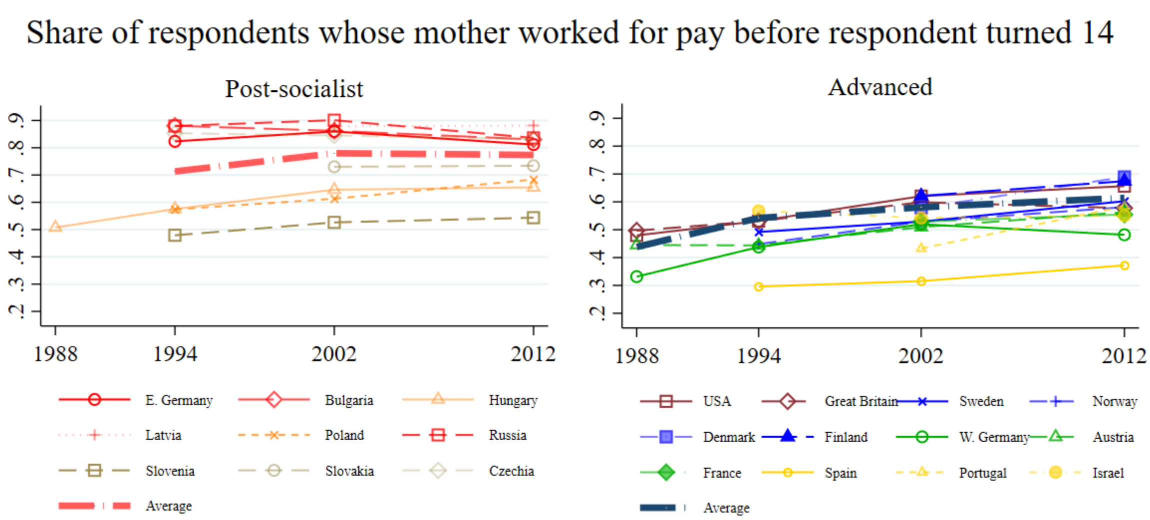


Figure 2.1: Retrospective evidence on female labor force participation rate prior to 1989 in post-socialist versus advanced economies

Notes: The graph presents averages for all respondents regardless of marital status and gender. The data for 1988 is available only for a few countries. Survey weights are applied if they are available; otherwise, unweighted statistics are reported.

To promote female employment, socialist governments had the orientation to alleviate the burden of household work for employed women through socialization of household work. For example, in a speech "Soviet Power and the Status of Women" in 1919 Lenin claimed that the Soviet Union made tremendous progress in terms of female emancipation and enlightenment and announced orientation for even more female empowerment in the future (Lenin, 1919).

While the Soviet Union improved the social position of women by providing a decent level of gender equality in the public sphere, gender equality in the private sphere was never reached.[3] When Stalin came to power in the late 1920s, rapid industrialization was prioritized over reduction of the burden of domestic work for women: women were actively pulled into employment in industry and agriculture with only limited childcare and educational facilities created to help with the transition (Goldman, 2001). The state created institutions of paid maternity leave and childcare that made it feasible, yet challenging for women to combine multiple roles and through propaganda praised women for their ability to smoothly navigate multiple responsibilities (Chatterjee, 2001).

The Soviet Union was not an exception in making women work "double shifts". For example, in East Germany, women received one paid day off for housework. If the state support was not sufficient for women to keep up with their responsibilities as workers and mothers, it was deemed their failure (Kranz, 2005). Men shared household responsibilities to some extent. According to some sources, men in East Germany participated in the household work more than men in West Germany (Geist, 2009). However, even though men were encouraged to help with household work, it was perceived as primarily a female domain.[4] Therefore, in most socialist countries "female emancipation" was reduced to increasing female employment rates without improving other dimensions of gender equality (LaFont, 2001).

Empirical evidence on household work division under socialism is rare. According to a few time use surveys, women devoted 2-2.5 times more hours to household work than men in the Soviet Union (Mishchenko, 2011). The micro data for such surveys are not publicly available nowadays. A woman's thoughts when answering such a survey and the double burden that Soviet women experienced due to combining work, motherhood, and home is depicted in a fictional novella "A week like any other" (Baranskaya and Lehrman, 1974). According to anecdotal evidence, many women were overwhelmed by the necessity of keeping up with their roles as workers, wives, and mothers (Funk and Mueller, 1993).

[3] The Soviet notion of "gender equality" at work often involved gender segregation. For example, in the 1930s, factory directors and coworkers massively resisted employing or working with women in skilled positions, so women ended up working at the lowest paid and, often, most physically demanding jobs. To make sure women were not prevented from working, the Soviet authorities started the process of classifying jobs into male and female. If the job could be done by a woman, it was classified as a "female job". "Male jobs" were those women typically could not do, and, therefore, superior to "female jobs". Such division enabled plants and factories to continue running while most men were at war (Goldman, 2001). Only in 2021 did Russia lift the early 1970s Soviet-era rule barring women from working in more than 350 professions that were considered harmful for reproductive health. An example of such professions is a truck driver or a boat captain (Maynes, 2021). Another mechanism frequently used to mitigate the discrimination from men in mixed enterprises was creating female-only brigades (Goldman, 2001).

[4] Ghodsee (2018) discusses state efforts to encourage men to participate in housework and childcare more actively in the 1950s in East Germany and Czechoslovakia.

Many of the cross-regime differences in female labor force participation disappeared by 1989-1991 when state-socialist countries started the transition to democracy. Figure 2.2 shows that survey participants from post-socialist and Western countries had quite similar female labor force participation rates of about 80%. In both groups of countries, married men of non-retirement age were more likely to be in the labor force and work for pay than married women, but the gender difference was smaller in post-socialist countries. These patterns are likely to reflect the complicated changes in the economic environment during the transition – multiple reforms, unraveling gender pay gaps, and a surge of unemployment rates in transition countries – that occurred in different years and with varying intensity (EBRD, 1995).[5]

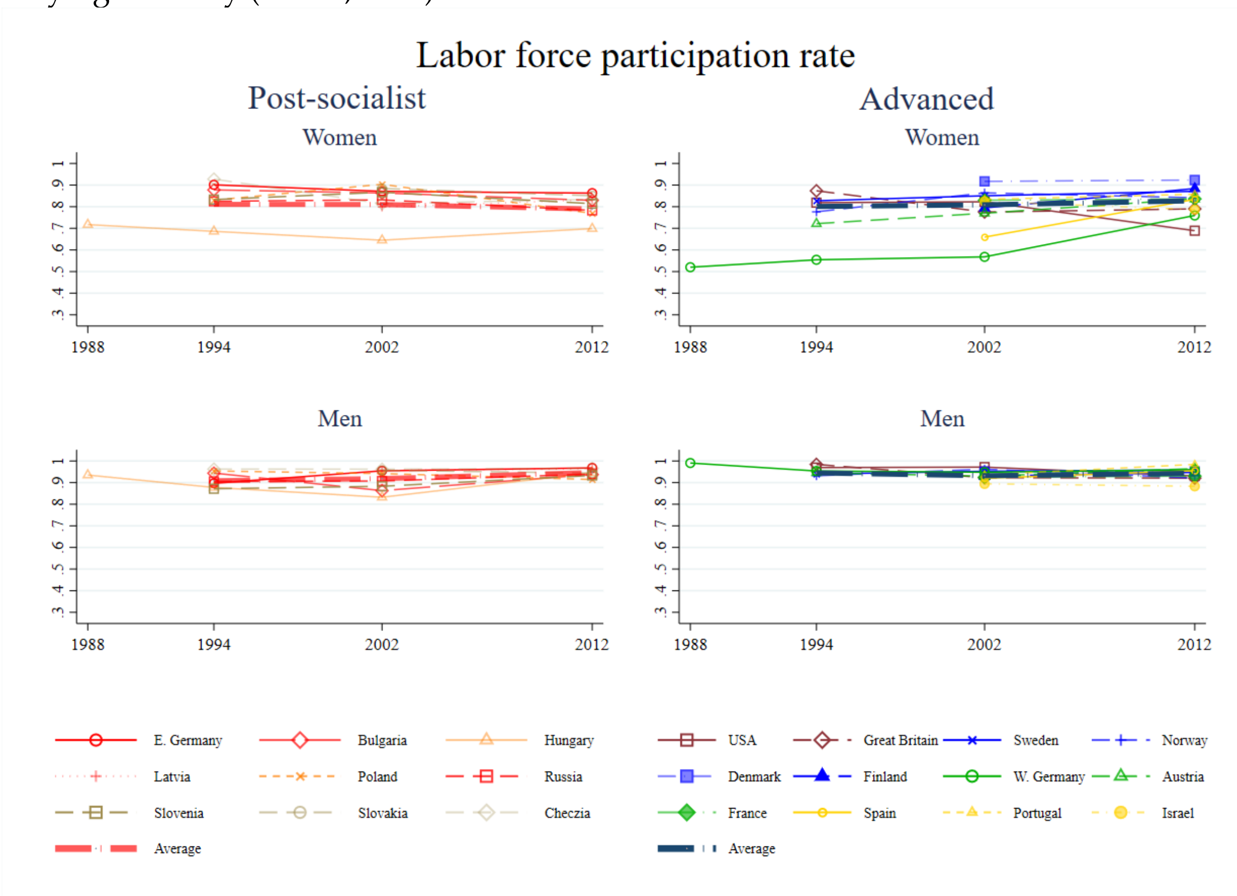


Figure 2.2: Labor force participation rates of married respondents

Notes: The graphs summarize averages for the sample of married respondents of non-retirement age. The data for 1988 is available only for two countries in a sample. Survey weights are applied if they are available; otherwise, unweighted statistics are reported.

While the democratic transition was unlikely to change the fundamental gender norms that became less traditional under socialism (Campa and Serafinelli, 2019), in many coun-

[5]Appendix Figures B.1 and B.2 illustrate the unemployment rates and gender pay gaps during the transition.

tries it coincided with remasculinization and orientation for women to return to the private sphere and traditional gender roles (Galligan et al., 2008). Rising unemployment rates and pay inequality in deregulated economies (Brainerd, 2000) contributed to forces pushing women into the domestic sphere. LaFont (2001) concludes that structural economic changes and the increase in power of the church and patriarchy altogether weakened women's position. However, in the end, how unpaid household work was organized in post-socialist countries during the transition period is an empirical question.

2.4 Data

The main dataset analyzed in this paper comes from the 1994, 2002, and 2012 waves of the "Family and Changing Gender Roles" module of the International Social Survey Programme (ISSP).[6] It is an annual program of cross-national collaboration made available by the Central Archive for Empirical Social Research of the University of Cologne (Scholz et al., 2014). The advantage of this data is that it contains information about hours devoted to different activities by married respondents and their spouses in a cross-section of countries including several post-socialist countries during the transition period.

The survey was designed using stratified random sampling. First, settlements were chosen, then households in these settlements were selected, and, finally, an adult household member was invited to participate in the interview. Examples of survey questions are provided in Appendix B.1. While there are many similarities across countries that enable cross-country comparison, there are differences in terms of question formulation, sample design, availability of weights, etc. For example, sampling weights are available only for some countries, and there is no total weight suitable for international comparisons. Therefore, by default, I do not use weights in the cross-country analysis. If I use weights for within-country summary statistics, I mention it in the notes to a table or figure.

The data from this survey module has been extensively studied in the literature (e.g., Fuwa, 2004; Ukhova, 2020). Many papers feature post-socialist countries in the sample, and some papers document how they stand out in cross-country comparisons. Among such peculiarities are low support of egalitarian gender norms, large number of hours women devote to household work, and relatively high male contributions to it (Treas and Tai, 2016; Baxter and Tai, 2016). Yet, evidence on how post-socialist countries compare to other welfare regimes in terms of time allocation is scarce.

Sample. I analyze trends in intra-household time allocation in five post-socialist countries (East Germany, Hungary, Latvia, Poland, Russia) and eleven advanced economies. The advanced economies represent four types of welfare regimes discussed by Kan et al.

[6] Detailed information about the data and country-specific details can be found in the codebooks: https://www.gesis.org/en/issp/modules/issp-modules-by-topic/family-and-changing-gender-roles. The next wave in this module is scheduled for 2022, and data will be published in spring/summer 2024.

(2011) – liberal (the USA), social democratic or Nordic (Denmark, Finland, Norway, Sweden), social capitalist or conservative (Austria, France, West Germany, Netherlands), and southern or Mediterranean (Israel, Italy, Spain). I use the terms welfare regimes and policy clusters interchangeably. Figure 3 illustrates how the clusters differ in terms of degree of state social support and traditionality of gender roles.

State support

Traditional gender roles

Mediterranean (Southern) ▪ Traditional gender ideology ▪ Less developed welfare policies than in the conservative cluster ▪ Higher concentration of women in part-time jobs and reliance on family-network welfare *Examples:* Israel, Italy, Spain	*Conservative (social capitalist)* ▪ Social insurance is used to improve the welfare ▪ Traditional gender ideology according to which family care is primarily women's responsibility dominates ▪ Welfare policies and labor market regulations are designed to assist women with family care *Examples:* Austria, France, W. Germany, Netherlands
Liberal: ▪ Limited state intervention in household welfare ▪ Reliance on market-based solutions (insurance, childcare) ▪ Women are both employed and mostly responsible for taking care of family *Examples:* the UK, the USA, Canada, Australia	*Nordic (social democratic):* ▪ Strong role of the state in provision of welfare policies ▪ State provides public childcare services and paternal leave ▪ Dual-earner families with egalitarian attitudes are prevalent in the country *Examples:* Denmark, Finland, Norway, Sweden

Figure 2.3: Welfare policy clusters (based on Kan et al. (2011))

Notes: State support is increasing from left to right and traditional gender roles are increasing from bottom to top.

The analysis sample in the paper is restricted to post-socialist and Western countries that a) participated in the survey at least in the 2002 and 2012 waves, when information about time allocation to household work by both spouses was collected, b) for which a welfare policy regime is known from prior literature, c) for which there is complete information about the key regression variables. These restrictions ensure that results are not driven by the changes in the composition of countries.[7] Since post-socialist countries are not mentioned in the conventional classification, I consider them as a separate group and discuss similarities and differences with the existing clusters. The number of observations from each country in the raw data and the final sample of married respondents used in the regression analysis are summarized in Appendix Table B.1. In the data, female respondents are slightly overrepresented, 40-60% of respondents are married, and the average respondent is 44-53 years old. Regression analysis is performed only for respondents who are 1) older than 18 years old and younger than 55-60 years old, and 2) married or having a steady partner with whom they reside at the time of the survey. I

[7] I excluded several countries from the sample because some key variables are missing for them. Specifically, information about hours worked by the spouse for the Czech Republic, Slovenia, and Slovak Republic is missing in 2002, and for Bulgaria and Great Britain it is missing in 2012. Results for these countries are reported in the Appendix B.4.

use the age cutoff of 55 years for women and 60 years for men as the likely lower bound for retirement age in the post-socialist countries. People who have reached the retirement age are likely to have more free time and therefore different routines which are beyond the scope of this paper. Restricting attention to married respondents reduces the sample by one-third to one-half in each country, but it is a necessary step to understand the behavior of couples. The respondents in the regression sample are heterogeneous in terms of education, number of children, and type of settlement they live in.

Creating Key Variables for Analysis. Regression analysis is performed for a couple (or a household) as a unit of observation. To test three theories of time allocation discussed in Section 2.2, I represent the key determinants according to these theories – income, time devoted to different activities (market work, household work, family care), and gender attitudes — from a female perspective. Because in the ISSP data the information about both respondent and spouse is reported by one respondent per household, I make several assumptions to infer the key variables from a female perspective. Specifically, I assume that i) all the marriages are mixed-gender[8], ii) household income consists of only the respondent's and spouse's income, i.e., income of all other family members is negligible[9], iii) spouses have the same gender attitudes. More details about processing the variables on time allocation are provided in Appendix B.2.

The average values of the key variables from a female perspective in the households with female and male respondents are summarized in the Appendix Table B.3. It allows the reader to judge the plausibility of the assumptions discussed above. With some discrepancies, the averages reported by male and female respondents are similar. However, in many countries, the share of household work and childcare performed by women in a couple's total time input is higher when the respondent is a woman. The larger discrepancy may imply that respondents may over- or underestimate how much time their spouses contribute to unpaid work at home. The gender attitudes of male and female married respondents within a country are similar, although female responses are generally more egalitarian. In the regression analysis, in addition to the all-gender sample, I also report results for a female-only sample that does not require additional assumptions.

2.5 Trends

This section discusses trends in household work allocation. I start from the extensive margin (who is usually engaged in a particular activity) and proceed with the intensive margin (how much time they devote to these activities).

[8]For the focal time period, most countries did not have legal same-sex marriages, and the overall share of couples that are not heterosexual couples is quite low in general (see Figure 1.1 from OECD report "Society at a Glance 2019").

[9]The facts that the average household size is close to 3 and couples have on average one child (Appendix Table B.1) support this assumption.

Extensive Margin. Table 2.1 summarizes who, a husband or wife, usually performs a particular household activity. It reports the average value of a categorical variable that takes a value of 1 (2) if all (most) of the work is done by a woman, 3 if it is split equally between spouses, and 4 (5) if most (all) work is done by a man. Similar to Fuwa (2004), the task-sharing index is the average value across all routine activities (i.e., excluding small repairs). The closer the values in the table are to 3, the more equally a particular task is allocated.

Table 2.1: Allocation of household activities in couples by policy cluster in 1994-2012

Country/year	Post-socialist (DE-E, HU, LV, PL, RU)			Liberal (US)			Nordic (DK, FI, NO, SE)			Conservative (AT, FR, DE)			Mediterranean (ES, IL, PT)		
	1994	2002	2012	1994	2002	2012	1994	2002	2012	1994	2002	2012	1994	2002	2012
Activities															
Cooking	2.22	1.98	1.95	2.38	2.43	2.35	2.37	2.39	2.58	2.31	2.01	2.15	.	2.01	2.17
Cleaning	.	2.08	1.97	.	2.32	2.19	.	2.33	2.40	.	2.04	2.08	.	2.06	2.12
Laundry	1.44	1.67	1.64	2.17	2.3	2.14	1.85	1.94	2.03	1.49	1.54	1.83	.	1.68	1.79
Groceries	2.35	2.45	2.45	2.42	2.4	2.35	2.59	2.63	2.68	2.44	2.39	2.47	.	2.50	2.71
Sick care	2.31	2.35	2.3	2.41	2.46	2.37	2.51	2.53	2.60	2.25	2.25	2.32	.	2.38	2.50
Small repair[a]	4.37	4.15	4.15	3.92	3.82	4.10	4.09	4.06	4.15	4.11	4.07	4.06	.	4.03	4.13
Task sharing	2.07	2.11	2.05	2.34	2.38	2.28	2.33	2.37	2.46	2.13	2.04	2.18	.	2.13	2.26
Sample size	2287	1323	977	410	297	322	1325	1763	1373	729	1159	1145	.	1025	1183

Notes: The table reports the average value of categorical variables taking value for who performs each activity always or mostly the woman (1-2), split equally or third person (3), always or mostly the man (4-5). [a] Task-sharing index is the average values across activities except for small repairs. The table reports data for the regression sample (married, non-retirement age, with key variables non-missing) in 2002 and 2012 and married respondents of non-retirement age in 1994.

The average value of a task-sharing index close to 2 indicates that in all the clusters, women are generally responsible for performing routine household work. Couples in post-socialist and conservative countries are less likely to share routine tasks than those in liberal and Nordic countries. Male partners in the sample are least likely to do laundry and most likely to share grocery shopping and care for sick household members. Cross-country differences in task allocation may reflect differences in availability of durable goods. For example, Table 2.1 shows that men are more likely to do laundry in the U.S. than in other policy clusters. This could be the case because doing laundry in communal laundry facilities in the U.S. is more pleasurable than doing laundry at home in other countries. Consistent with prior literature, the only task that men are much more likely to engage in than women is small repairs which is typically referred as nonroutine household work because it needs to be done infrequently and usually can be postponed.

There are different trends in task-sharing in the policy clusters. The inverse U-shaped pattern is observed in post-socialist and liberal policy clusters where after a slight increase in 1994-2002, task sharing decreased again in 2012. At the same time, Nordic, conservative, and Mediterranean countries experienced an increase in task sharing during 1994-2012. The allocation of household work at the extensive margin on the country level is summarized in Appendix Figure B.4.

Intensive Margin. The amount of time married women devote to different activities in absolute and relative terms by policy cluster is summarized in Table 2.2. Unlike task sharing, the questions about hours devoted by spouses to unpaid work were not included in the survey in 1994. Therefore, the intensive margin analysis is restricted to 2002-2012. There is substantial variation in time allocation between post-socialist and advanced economies as well as within the post-socialist group. The average number of hours devoted by women to market work varies between 20 and 30 and is the largest in post-socialist and Nordic countries. In relative terms, female input accounts for 35-to-45% of the total time spouses devote to market work.

Table 2.2: Married women's time allocation by policy cluster in 2002-2012

Country/year	Post-socialist (DE-E, HU, LV, PL, RU)		Liberal (US)		Nordic (DK, FI, NO, SE)		Conservative (AT, FR, DE-W)		Mediterranean (ES, IL, PT)	
	2002	2012	2002	2012	2002	2012	2002	2012	2002	2012
Market work										
Abs., hours	31.85	29.05	29.28	24.53	30.77	33.11	24.83	25.43	26.34	25.44
Gap, hours	-9.56	-10.23	-14.33	-16.13	-8.85	-6.25	-14.75	-13.38	-13.94	-12.69
Rel., share	0.424	0.409	0.383	0.349	0.426	0.451	0.368	0.38	0.376	0.393
Household work										
Abs., hours	22.11	24.18	13.3	16.91	12.14	11.73	17.00	15.33	22.27	24.02
Gap, hours	10.49	11.03	7.08	8.73	6.01	4.35	11.75	9.39	15.56	15.06
Rel., share	0.668	0.667	0.67	0.655	0.67	0.622	0.757	0.713	0.751	0.719
Family care										
Abs., hours	.	19.65	.	34.09	.	18.25	.	20.11	.	23.13
Gap, hours	.	9.79	.	14.67	.	5.6	.	9.66	.	10.18
Rel., share	.	0.633	.	0.63	.	0.577	.	0.629	.	0.618
Income share	0.42	0.40	0.34	0.36	0.40	0.43	0.36	0.36	0.36	0.39
Sample size	1324	983	297	322	1763	1378	1161	1153	1025	1198

Notes: The table summarizes time devoted to a particular activity as hours per week. The absolute hours correspond to raw data subject to the "upper bound" imputation procedure described in Online Appendix B. The gap denotes the average value of the difference between female and male hours in a couple. Relative hours are the average value of the ratio of hours devoted to a particular activity by a female divided by the household total (sum of male and female input). The summary statistics are reported for respondents in the regression sample (married, non-retirement age, with key variables non-missing). Survey weights are not applied.

Household work on average requires 12-25 hours of woman's time per week which constitutes 2/3-to-3/4 of the household total. The average relative female input to household work remained unchanged between 2002 and 2012 in post-socialist countries and declined in all other policy clusters. In absolute terms, women from post-socialist and Mediterranean countries devote the largest number of hours to household work. The absolute time input to household work reflects many factors, including availability of time-saving household appliances. The socialist regimes were notorious for the scarcity of such goods which substantially contributed to the double burden, thus the pattern observed today may be echoing the past.

The information about the time devoted to care for children, sick, and elderly family members is available only in 2012. Family care takes up on average 18-35 hours per week of woman's time which constitutes about 60% of the total household time input. In Nordic countries, relative female input to family care is the smallest. Given the male parental leave that is gaining popularity there, it is not surprising that Nordic countries stand out in terms of equality of time allocation to family care (Almqvist and Duvander, 2014). Women in post-socialist countries also contribute relatively little time to family care, presumably due to kindergartens and other facilities which many post-socialist countries inherited from the pre-1990s period.

Table 2.2 highlights that the patterns of time allocation in the post-socialist countries are a mix of those observed in different policy clusters. Women from post-socialist countries are "leading" in terms of hours they devote to household work and market work. In Nordic countries, where women work for pay about the same number of hours as in post-socialist countries, they spend half as much time on household work and about the same amount of time on family care. In Mediterranean countries, women devote about the same amount of time to household work, but they also work fewer hours for pay and devote a few hours more to family care. In conservative countries, women devote about the same amount of time to family care, but less time to both household work and market work. In the liberal policy cluster, women devote much more time to family care but less time to household work and market work. Despite differences in levels, in relative terms, different activities are allocated about as well in post-socialist countries as in other policy clusters. However, such "equality" in relative terms is accompanied by gender segregation in tasks, as shown in Table 2.1. There are also differences in trends by policy clusters.

The post-socialist countries appear to diverge from the global trend of equalization of relative input to household work by women and men. In post-socialist countries, female relative input in household work has practically not changed between 2002 and 2012. Relative women's input to market work and share of income contributed by women has slightly declined in post-socialist countries and increased in other policy clusters. Couples in the post-socialist countries were less likely to share the tasks traditionally done by women in 2012 than in 2002. Women on average contribute 30-40% of household income and this share has been increasing in all groups of countries except for post-socialist ones. The allocation of household work at the intensive margin on the country level is summarized in Appendix Table B.4.

Gender Attitudes. I consider gender attitudes as a proxy for gender perspective, one of the key theoretical determinants of time allocation. They are summarized in Table 2.3. Respondents in post-socialist countries appear to approve of the traditional division of gender roles. Over 40% of respondents in post-socialist countries, twice as many as in the advanced economies, agree with traditional division of gender roles (i.e., that men's job is to earn money and women's job is looking after home and family). At the same time, over 70% of married people in post-socialist countries agree that both husband and wife should contribute to household income. Moreover, about 1/2-to-2/3 of the respondents agreeing with traditional division of gender roles in post-socialist countries also agree that both spouses should contribute to income. This pattern suggests that the burden of working “double shifts” has become a norm for many households in post-socialist countries. It could also rationalize why segregation into male and female activities persists as suggested by the trend in the task-sharing index.

Table 2.3: Gender attitudes of married respondents by country group in 1994-2012

Country/year	Post-socialist (DE-E, HU, LV, PL, RU)			Liberal (US)			Nordic (DK, FI, NO, SE)			Conservative (AT, FR, DE)			Mediterranean (ES, IL, PT)		
	1994	2002	2012	1994	2002	2012	1994	2002	2012	1994	2002	2012	1994	2002	2012
=1 if trad.div	0.49	0.41	0.42	0.17	0.19	0.20	0.09	0.06	0.05	0.30	0.18	0.14	.	0.20	0.12
=1 if both work	0.69	0.8	0.79	0.49	0.54	0.54	0.66	0.72	0.79	0.66	0.71	0.73	.	0.88	0.91
= 1 if trad&work	0.28	0.28	0.28	0.06	0.09	0.08	0.03	0.03	0.03	0.16	0.10	0.08	.	0.15	0.10
=1 if housewife	0.55	0.38	0.42	0.53	0.58	0.61	0.27	0.27	0.27	0.37	0.33	0.29	.	0.32	0.31
Sample size	2323	1324	983	425	297	322	1333	1763	1378	734	1161	1153	.	1025	1198

Notes: The table reports the share of respondents who either agree or strongly agree with the following claims: 1) “Men’s job is to earn money and women’s job is to look after the home and family”; 2) “Both husband and wife should contribute to household income”, 3) statements 1 & 2, 4) “Being a housewife is as rewarding as working for pay”. The summary statistics are reported for respondents in the regression sample (married, non-retirement age, with key variables non-missing). Survey weights are not applied.

Democratic transition has unlocked a career option that was not readily available for many women under socialism — being a housewife. The share of respondents agreeing that “Being a housewife is just as fulfilling as working for pay” is among the highest in post-socialist countries. The U-shaped pattern may indicate learning the pros and cons of being a housewife over time. According to Appendix Table B.4, women generally have more egalitarian gender attitudes than men. However, there is substantial heterogeneity within post-socialist countries block. East Germany stands out as the country with much more egalitarian gender attitudes. It is also the country with the largest task-sharing index in the post-socialist countries, and the smallest number of hours women devote to household work and family care. It suggests that the East German experience is not representative of the experience of other post-socialist countries. It is an important finding given that the East versus West Germany case is the most studied setting in the literature about the effects of socialism.

Summary. Overall, in all countries, women perform more of the unpaid household work both in terms of extensive and intensive margins and contribute less than half of household income. In the post-socialist countries, women devote the longest hours to household work but about the same proportion as their western counterparts. Post-socialist countries also have more traditional gender attitudes and more gender segregation of household work relative to advanced economies.

Next, I perform the regression analysis to understand what factors rationalize the differences in time allocation across policy regimes.

2.6 Methods

Regression Specification. My main regression specification is motivated by three theories of household work division in Section 2.2. This methodology is widely used in prior literature for different countries and datasets (e.g., Bianchi et al., 2000; Fuwa, 2004).

The unit of analysis is a household h residing in a country c consisting of a married couple, and possibly other household members. The question is how female f and male m spouses in the household h allocate unpaid work at home conditional on their employment, individual, and household characteristics. I analyze the determinants of time allocation using the regression equation as follows.

$$\widetilde{h_{fht}^{home}} = \alpha_0 + \alpha_1 PS_c + \boldsymbol{TA}_{fhc}\, \boldsymbol{\beta}_1 + \boldsymbol{BP}_{fhc}\, \boldsymbol{\beta}_2 + \boldsymbol{GP}_{fhc}\, \boldsymbol{\beta}_3 + \boldsymbol{X}_{hr}\boldsymbol{\theta} + \epsilon_{fhc} \tag{2.1}$$

The outcome variable is the share of weekly time devoted to the household work by a woman in household total, i.e. $\widetilde{h_{fht}^{home}} = h_{fht}^{home}/(h_{fht}^{home} + h_{mht}^{home})$. The explanatory variables are suggested by the theories: $\boldsymbol{TA}$ denotes time availability and includes share of time devoted by a woman to work for pay in household total (i.e., $\widetilde{h_{fht}^{mkt}} = h_{fht}^{mkt}/(h_{fht}^{mkt} + h_{mht}^{mkt})$) and number of children, $\boldsymbol{BP}$ denotes bargaining power (or relative resources) and includes a woman's share in total household income and education categorical variable, and $\boldsymbol{GP}$ denotes gender perspective and includes dummies reflecting attitudes to gender equality. $PS = \mathbb{1}\{c \in PostSoc\}$ is an indicator variable for a post-socialist country. To understand how the role of the key determinants differs in post-socialist countries relative to advanced economies, I introduce an interaction term of each determinant with the post-socialist dummy. Control variables $\boldsymbol{X_{hc}}$ include respondent's education dummy, number of children, religion, urban-rural status, logarithm of family income, cohort fixed effect, and year fixed effects (for household work). The coefficients of interest are the ones on the post-socialist dummy and its interaction with the key determinants of time allocation.

For each regression specification, I separately analyze time allocation to household work and family care. For each outcome variable, I estimate the results for two subsamples – all married respondents and female-only married respondents. The estimates for the female-only subsample do not rely on the assumptions necessary to infer female input to household total in the households with male respondents. Since the sample size is rather small and the cross-country data cannot be plausibly viewed as independent draws from a common distribution, I bootstrap standard errors clustered on the country level (with 500 repetitions).

Decomposition. To better understand the contribution of different factors to intrahousehold time allocation in the post-socialist countries (group S) relative to advanced economies (group A), I perform a Kitagawa-Blinder-Oaxaca decomposition of the difference in the average share of unpaid household work performed by women in these countries $E(Y_A) - E(Y_S)$:

$$\mathrm{R} = \mathrm{E}(\mathrm{Y}_A) - E(Y_S) = \underbrace{[E(X_A) - E(X_S)]'\beta_S}_{\text{Endowment effect}} + \underbrace{E(X_S)'(\beta_A - \beta_S)}_{\text{Coefficient effect}} + \underbrace{[E(X_A) - E(X_S)]'(\beta_A - \beta_S)}_{\text{Interaction effect}}$$

The decomposition shows what value the difference in outcomes would take 1) if households in the post-socialist countries had the same observable characteristics as the households in advanced economies (endowment effect, or the effect of observables), 2) if the contribution of the different factors to intra-household time allocation in the post-socialist countries was the same as in the advanced economies (coefficient effect or "price effect"), and 3) the residual (interaction effect). Next, I discuss the association between female input to unpaid work at home and theoretical determinants of time allocation.

2.7 Regression Results

Baseline Specification. The regression results about gender time allocation along the intensive margin in post-socialist countries relative to the advanced economies are reported in Tables 2.4 and 2.5. The estimates for the baseline specification reported in columns 1 and 5 of Table 2.4 indicate that time allocation to household work is broadly consistent with theoretical predictions. Female relative input to market work and to household income are negatively associated with the share of time a woman devotes to household work with point estimates respectively -0.15 and -0.06 in the all-gender sample, both significant at 0.01 level. This result holds in the female-only sample as well. Approval of traditional gender roles is positively associated with increase of share of household work performed by women (point estimate is 0.02 in the all-gender sample). Agreement with the claim that both spouses should contribute to household income is negatively associated with female input to household work (point estimate is -0.02 in the all-gender sample).

The baseline results for time allocation to family care are similar except for a smaller and statistically insignificant coefficient for the share of income provided by women (Table 2.5). While relative woman's input to market work is negatively associated with her relative input to family care (point estimate -0.13, statistically significant at 0.01 level), women who earn more do not tend to perform a significantly smaller share of family care (point estimate is close to 0, statistically insignificant). This may be explained by the different nature of care activities, many of which could be more pleasant than doing household chores.

Interaction Terms. I include an interaction of each of the key determinants with the post-socialist indicator variable to test whether different determinants contribute more or less to time allocation in the post-socialist countries relative to the advanced economies. The coefficients on interaction terms, one at a time, are reported in columns 2-4 and 6-8 of Tables 2.4 and 2.5. For the share of time devoted to household work in Table 2.4, the coefficient on interaction term with relative female input to market work is positive and significant (0.08 in the all-gender sample). It indicates that time allocation in post-socialist countries is less sensitive to female time availability. Given that the coefficient on the interaction term of share of income contributed by a woman with the post-socialist dummy is small and statistically insignificant, this factor appears to be equally important for time allocation in post-socialist countries and advanced economies. Similarly, agreeing that both a man and a woman should contribute to household income is equally important in both groups of countries. At the same time, the negative interaction term with traditional gender attitudes suggests that the positive association in the baseline specification appears to be almost entirely driven by the advanced economies.

Table 2.4: OLS regression results for determinants of time allocation by married women to household work in 2002-2012

	Panel A: All respondents				Panel B: Female respondents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widetilde{h^{fem}_{mkt}}$	-0.15***	-0.17***	-0.15***	-0.15***	-0.15***	-0.17***	-0.15***	-0.15***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\widetilde{h^{fem}_{mkt}} \times PS$		0.08***				0.06***		
		(0.02)				(0.02)		
$\widetilde{y^{fem}}$	-0.06***	-0.06***	-0.06***	-0.06***	-0.04***	-0.04***	-0.05***	-0.04***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
$\widetilde{y^{fem}} \times PS$			0.02				0.03	
			(0.02)				(0.03)	
=1 if traditional division	0.02***	0.02***	0.02***	0.03***	0.02***	0.02***	0.02***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
trad.division × PS				-0.04***				-0.04***
				(0.01)				(0.01)
=1 if both should work	-0.02***	-0.02***	-0.02***	-0.01**	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
both work × PS				-0.01				-0.01
				(0.01)				(0.01)
=1 if 2012	-0.03***	-0.03***	-0.03***	-0.03***	-0.02***	-0.02***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
=1 if post-socialist (PS)	-0.06*	-0.09***	-0.06*	-0.03	-0.07***	-0.10***	-0.08***	-0.06**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
PS × 2012	0.04***	0.04***	0.04***	0.04***	0.05***	0.05***	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	10,604	10,604	10,604	10,604	5,454	5,454	5,454	5,454
R-squared	0.11	0.12	0.11	0.11	0.12	0.12	0.12	0.12
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The outcome variable is the share of time a woman in a couple devotes to activity mentioned in the part name. Standard errors clustered on country level and estimated using bootstrap are in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. No survey weights are applied. The descriptive statistics for the analysis sample are provided in Table 2.2 and 2.3. Notation: "PS" – indicator variable for a post-socialist country. $\widetilde{h^{fem}_{mkt}}$ – share of time devoted by a female household member to work for pay in household total. y^{fem} – share of female income in household income. "=1 if traditional division" – respondent agrees with a claim that "Men's job is to earn money and women's job is to look after home and family". "=1 if both should work" respondent agrees with a claim that "Both husband and wife should contribute to household income". Control variables include respondent's education dummy, number of children, cohort fixed effect, religion, urban-rural status, and logarithm of family income, year fixed effects (for household work). Unabridged regression results are reported in Table B.5 and Table B.6.

For the time allocation to family care, most factors, including time availability, appear to be as important in the post-socialist countries as in the advanced economies (Table 2.5). The negative association between agreeing with the claim that both spouses should contribute to the household income and that women's contribution to family care is statistically significant only for the post-socialist countries.

Table 2.5: OLS regression results for determinants of time allocation by married women to family care in 2012

	Panel A: All respondents				Panel B: Female respondents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\widetilde{h^{fem}_{mkt}}$	-0.13***	-0.13***	-0.13***	-0.13***	-0.14***	-0.15***	-0.14***	-0.14***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
$\widetilde{h^{fem}_{mkt}} \times PS$		0.02				0.05		
		(0.04)				(0.06)		
$\widetilde{y^{fem}_{mkt}}$	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\widetilde{y^{fem}_{mkt}} \times PS$			-0.02				-0.02	
			(0.04)				(0.05)	
=1 if traditional division	0.02*	0.02*	0.02*	0.02*	0.02	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
trad.division × PS				-0.01				0.00
				(0.02)				(0.04)
=1 if both should work	-0.02**	-0.02**	-0.02**	-0.01	-0.01	-0.01	-0.01	0
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
both work × PS				-0.04***				-0.04*
				(0.02)				(0.02)
=1 if post-socialist (PS)	0.02	0.01	0.02	0.05***	0.01	-0.01	0.02	0.05
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)
Observations	4,829	4,829	4,829	4,829	2,542	2,542	2,542	2,542
R-squared	0.06	0.06	0.06	0.06	0.08	0.08	0.08	0.08
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The outcome variable is the share of time a woman in a couple devotes to activity mentioned in the part name. Standard errors clustered on country level and estimated using bootstrap are in parentheses: *** p<0.01, ** p<0.05, * p<0.1. No survey weights are applied. The descriptive statistics for the analysis sample are provided in Table 2.2 and 2.3. Notation: "PS" – indicator variable for a post-socialist country. $\widetilde{h^{fem}_{mkt}}$ – share of time devoted by a female household member to work for pay in household total. y^{fem} – share of female income in household income. "=1 if traditional division" – respondent agrees with a claim that "Men's job is to earn money and women's job is to look after home and family". "=1 if both should work" respondent agrees with a claim that "Both husband and wife should contribute to household income". Control variables include respondent's education dummy, number of children, cohort fixed effect, religion, urban-rural status, and logarithm of family income, year fixed effects (for household work). Unabridged regression results are reported in Table B.5 and Table B.6.

Policy Clusters. To compare the contribution of each factor within the policy cluster, I estimate baseline specification separately for each cluster in Table 2.6. The results show that while time availability is quantitatively the most important predictor, it plays a more important role in division of household work in liberal, Mediterranean, and Nordic countries than in post-socialist and conservative clusters. For female income share, it is about as important as time availability in post-socialist countries, but it plays a smaller role than time availability in other policy clusters. Unlike other policy clusters, the association between female relative input to household work and traditional gender attitudes is weaker compared to the association with the agreement that both spouses should contribute to household income. Female time availability is also a less important predictor of time allocation to family care in post-socialist countries than in liberal and Mediterranean clusters. Female income share has a significant effect on time allocation to family care only in the Nordic policy cluster. Agreeing that both respondents should contribute to household income is associated with a lower share of time women spend on family care only in the post-socialist and Nordic cluster.

Table 2.6: OLS regression results for determinants of time allocation by married women in 2002-2012 by policy cluster

	Panel A: All respondents					Panel B: Female respondents				
	Post-socialist	Liberal	Nordic	Conser-vative	Mediter-ranean	Post-socialist	Liberal	Nordic	Conser-vative	Mediter-vative
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Part 1: Household work, 2002-2012										
$\widetilde{h_{mkt}^{fem}}$	-0.08	-0.21	-0.17	-0.13	-0.19	-0.10	-0.15	-0.19	-0.10	-0.17
	(0.02)***	(0.04)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.05)***	(0.02)***	(0.03)***	(0.03)***
$\widetilde{y^{fem}}$	-0.08	-0.04	-0.08	-0.05	-0.03	-0.02	-0.03	-0.05	-0.10	-0.03
	(0.02)***	(0.03)	(0.02)***	(0.02)**	(0.02)	(0.03)	(0.05)	(0.03)*	(0.03)***	(0.02)
=1 if traditional division	0.00	0.03	0.05	0.02	0.02	0.01	0.04	0.05	0.03	0.04
	(0.01)	(0.02)	(0.02)***	(0.01)**	(0.01)*	(0.01)	(0.03)	(0.03)*	(0.01)**	(0.02)**
=1 if both should work	-0.03	-0.01	-0.01	-0.02	-0.03	-0.01	0	-0.01	-0.01	-0.03
	(0.01)***	(0.02)	(0.01)*	(0.01)***	(0.01)***	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
=1 if 2012	-0.01	0	-0.04	-0.03	-0.02	0.01	-0.01	-0.02	-0.02	-0.03
	(0.01)*	(0.02)	(0.01)***	(0.01)***	(0.01)*	(0.01)	(0.03)	(0.01)**	(0.01)**	(0.01)*
Observations	2,307	619	3,141	2,314	2,223	1,172	324	1,572	1,302	1,084
R-squared	0.07	0.13	0.12	0.11	0.13	0.07	0.12	0.12	0.12	0.12
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Part 2: Family care, 2012										
$\widetilde{h_{mkt}^{fem}}$	-0.08	-0.31	-0.09	-0.09	-0.15	-0.07	-0.33	-0.05	-0.17	-0.16
	(0.03)**	(0.07)***	(0.03)***	(0.03)***	(0.02)***	(0.05)	(0.10)***	(0.04)	(0.04)***	(0.03)***
$\widetilde{y^{fem}}$	-0.03	-0.04	0.07	-0.02	0.03	-0.05	-0.06	0.08	0	0.02
	(0.03)	(0.07)	(0.03)**	(0.03)	(0.02)	(0.05)	(0.10)	(0.04)*	(0.05)	(0.04)
=1 if traditional division	0.01	-0.02	0.03	0.02	0.02	0.02	-0.02	0.04	0.01	0.02
	(0.01)	(0.04)	(0.03)	(0.02)	(0.02)	(0.02)	(0.06)	(0.04)	(0.02)	(0.03)
=1 if both should work	-0.05	0.05	-0.02	0	-0.02	-0.04	-0.01	-0.02	0.01	-0.03
	(0.01)***	(0.03)	(0.01)*	(0.01)	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)	(0.03)
Observations	882	322	1,344	1,107	1,174	472	161	666	631	612
R-squared	0.08	0.14	0.04	0.07	0.10	0.09	0.30	0.05	0.13	0.10
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The outcome variable is the share of time a woman in a couple devotes to activity mentioned in the part name. Standard errors estimated using bootstrap are in parentheses: *** p<0.01, ** p<0.05, * p<0.1. No weights are applied. The descriptive statistics for the analysis sample are provided in Table 2.2 and 2.3. Notation: $\widetilde{h_{mkt}^{fem}}$ – share of time devoted by a female household member to work for pay in household total. $\widetilde{y}^{fem}$ –- share of female income in household income. "=1 if traditional division" – respondent agrees with a claim that "Men's job is to earn money and women's job is to look after home and family". "=1 if both should work" means that respondent agrees with a claim that "Both husband and wife should contribute to household income". Other control variables include year fixed effect, respondent's education dummy, number of children, cohort fixed effect, religion, urban-rural status, and logarithm of family income.

Trends. The coefficient on the post-socialist dummy in Table 2.4 (-0.06 in the all-gender sample) indicates that in 2002 the average relative female input to household work was smaller in the post-socialist relative to advanced economies after controlling for all other variables. In 2012, it decreased by less in post-socialist countries than in advanced economies. This pattern can also be noticed in Table 2.6. It suggests that differences in sample composition have some effect on the trends discussed in Section 2.5.

Aggregate Measures of Inequality. Existing literature documents a positive association between aggregate measures of gender inequality and inequality in division in household work (e.g., Fuwa, 2004). To examine the role of aggregate measures of inequality, I added the country-level gender unemployment gap and gender pay gap in hourly earnings from the United Nations Economic Commission for Europe (UNECE) database to the baseline regression specification. I consider them as aggregate analogs for time availability and relative resources. The results yield small and insignificant coefficients on aggregate measures of inequality suggesting that these factors are unlikely to be important predictors of intra-household time allocation after controlling for individual- and household-level factors. Aggregate factors also have weak explanatory power at explaining differences across policy clusters which may be due to limited variation because all respondents from the same country have the same value assigned in a given year.

Female-Only Sample. Although there are some discrepancies in the results in the all respondents sample (panels A) and female only sample (panel B) of Tables 2.4-2.6 , it is reassuring that the main results are qualitatively similar across samples. It suggests that assumptions to increase the sample size discussed in Section 2.4 are not pivotal for the regression results.

Decomposition Results. Next, I apply the Kitagawa-Blinder-Oaxaca decomposition method to better understand the source of differences across policy clusters unexplained by the conventional determinants of time allocation. Table 2.7 reports the results for post-socialist countries relative to the Nordic policy cluster as a benchmark for egalitarian division. The decompositions for other clusters are reported in Appendix Table B.7.

The difference in means of relative female input to household work is close to 0 in 2002, but negative in 2012 because relative female input decreased substantially in Nordic countries and decreased a little (or even increased in the female-only sample) in post-socialist countries. The endowment effect shows that if women from post-socialist countries had the same characteristics as in Nordic countries, they would be doing a relatively smaller share of household work. The endowment effect is mostly driven by differences in demographic characteristics and not by differences in the key determinants considered in Section 2.6. In 2012, the endowment effect is positive but smaller than in 2002 and associated with relatively large standard errors, which implies that there were fewer differences in characteristics in the two samples that could unambiguously explain differences in relative female time input to household work.

The coefficient effect is largely imprecisely estimated and does not allow for drawing strong conclusions about the role of differences in returns to controlled characteristics in explaining the observed differences in time allocation. Qualitatively, the negative sign of the coefficient effect in 2012 suggests that differences in the importance of controlled characteristics in 2012 mitigate the differences in the share of household work performed by women in two policy clusters. The interaction effect, which captures the differences explained neither by differences in analyzed factors nor by returns to them, is consistently negative and often quite sizeable but imprecisely estimated. The direction of the effect suggests that endowment and coefficient effects together tend to overexplain the differences across policy clusters.

Table 2.7: Kitagawa-Blinder-Oaxaca decomposition of share of time devoted by married women to different activities in post-socialist countries relative to the Nordic policy cluster in 2002-2012

	Panel A: All respondents			Panel B: Female respondents		
	Household work, 2002	Household work, 2012	Family care, 2012	Household work, 2002	Household work, 2012	Family care, 2012
Mean, Nordic	0.671	0.607	0.577	0.697	0.661	0.627
Mean, post-socialist	0.67	0.66	0.634	0.678	0.72	0.653
Difference in means	**0.001**	**-0.053**	**-0.057**	**0.02**	**-0.058**	**-0.026**
	(.022)	(.016)	(.016)	(.022)	(.03)	(.037)
Endowment effect	**0.072**	**0.035**	**0.023**	**0.046**	**0.018**	**-0.003**
	(.025)	(.023)	(.03)	(.024)	(.042)	(.048)
Determinants	*0.002*	*-0.005*	*-0.009*	*0.011*	*-0.019*	*-0.018*
	(.004)	*(.003)*	*(.006)*	*(.009)*	*(.004)*	*(.012)*
Coefficient effect	**-0.029**	**-0.056**	**-0.008**	**0.035**	**-0.063**	**0.019**
	(.026)	(.042)	(.023)	(.027)	(.045)	(.037)
Determinants	*-0.035*	*0.012*	*0.069*	*-0.06*	*-0.034*	*0.071*
	(.024)	*(.019)*	*(.032)*	*(.044)*	*(.025)*	*(.04)*
Interaction effect	**-0.042**	**-0.032**	**-0.072**	**-0.061**	**-0.014**	**-0.042**
	(.028)	(.044)	(.036)	(.021)	(.051)	(.055)

Notes: The regression equation used as a basis for decomposition includes all the variables used in column 1 of Appendix Table B.5 and B.6. Determinants stand for relative female input to market work, household income share contributed by women, and indicator variables for gender attitudes. The decomposition was performed using the oaxaca module in Stata (Jann, 2008).

As to family care, in 2012, women in post-socialist countries performed a larger share of family care relative to women from the Nordic policy cluster. This difference can only partially be explained by differences in key determinants (endowment effect). A positive coefficient effect for the key determinants suggests that had women in post-socialist countries been assigned the same coefficients on their relative input to market work, income share, and gender attitudes as those of women from Nordic countries, they would be performing a lower share of family care. The total coefficient effect is also positive but imprecisely estimated. A large interaction effect that exceeds the differences in means, although imprecisely estimated, implies that some additional factors or unobservable characteristics might explain why the allocation of household work is different between policy clusters.

Summary. According to regression analysis, conventional determinants of time allocation are significant predictors for couples in post-socialist countries. Time allocation to household work there is most sensitive to time availability. However, the sensitivity to this determinant is smaller than in other policy clusters. An increase in relative female income is associated with a decrease in female input to household work by about the same amount in post-socialist countries and advanced economies. At the same time, traditional gender attitudes in post-socialist countries appear a less important predictor of household work allocation relative to advanced economies. This could be an indicator of reduced attention to gender norms when allocating work due to exposure to contradictory state ideology for decades. The decomposition results help to explain whether changes in different characteristics would increase or decrease inequality in time allocation to unpaid work at home. At the same time, they do not allow drawing unambiguous conclusions about what explains the differences and similarities in time allocation in different policy clusters.

2.8 Discussion and Conclusions

State-socialist countries are viewed by many as the ones providing outstanding support for female emancipation and gender equality (Ghodsee and Mead, 2018). This paper focuses on the aspect at which the socialist countries were presumably backward — gender equality in private life and, in particular, division of household work. I compare intra-household time allocation to unpaid work at home by married couples in several post-socialist and western economies after the fall of communism in 1989. I use data from the "Family and Changing Gender Roles" module of the International Social Survey Programme.

Has gender equality in the private sphere improved during the transition from socialism to a market economy? Have the households stopped following socialist-era rules and customs? My analysis suggests negative answers to these questions. Despite a lot of similarities in household work division patterns across regimes during the transition period, some differences persist.

On the one hand, there is convergence in labor force participation rates and not so drastic differences in measures of inequality of household work. On the other hand, time allocation to household work in post-socialist countries is less sensitive to women's relative time availability, a conventional determinant of time allocation according to the theories of time allocation. In line with Fuwa (2004), the slope coefficient from regression of relative female input to household work on relative female input to work for pay in post-socialist countries is 1/2-2/3 of the coefficient in the advanced economies. This could be viewed as evidence of persistence of the norm of working "double shifts" for women throughout the early transition period.

Several limitations need to be kept in mind when interpreting the findings. First, regression results summarize an association between time allocation and its potential determinants and should not be interpreted as causal. In fact, the causality between time allocation to household work and female income share could go in the opposite direction (Albanesi and Olivetti, 2009). Second, the results are obtained using a cross-country survey with limited sample sizes. This is a common problem in the literature because very few surveys are suitable for the analysis of intra-household time allocation in post-socialist countries. I mitigate it by performing robustness checks and reporting bootstrap standard errors. Third, measurement errors and misreporting are possible. For example, answers could be biased because the same person answered questions about contribution of both spouses to different activities. Section 4 discusses assumptions and data cleaning procedures for computing relative female input to household activities and income. I report summary statistics by gender of the respondent and perform robustness checks to ensure that the results are not driven by my assumptions rather than patterns in the data.

Fourth, while my analysis focuses on comparing the post-socialist countries with other policy clusters, there may be substantial heterogeneity within the post-socialist cluster that I do not discuss. LaFont (2001) mentions that the Soviet bloc and East Germany experienced different forms of socialism. Descriptive statistics in my paper also support this conclusion. It also suggests that patterns in Russia may differ from patterns of other former members of the Soviet Bloc (Bystydzienski, 1989). Additionally, my results should not be extended to non-European post-socialist countries which are not represented in the data analyzed and differ from other post-Soviet countries in terms of religion and culture. Finally, I document that conventional determinants of intra-household time allocation are less important in post-socialist countries and that cross-cluster differences are driven mostly by unobservables, i.e. factors I do not control for. Examination of the contribution of such determinants is a task for future research.

There are several promising candidates for such unobserved determinants. First, heterogeneity behind the policy clusters in terms of how burdensome household work is and differences in gender ideologies. The burden of household work includes access and feasibility of state-provided and private services (e.g., cleaning services, childcare) as well as affordability of household durables. These factors vary across policy regimes, but they also vary across households. While there is no data about durables availability for a large cross-section of post-socialist countries, they may be available for a smaller sample. Work hours flexibility and the size of reward for inflexible hours are additional factors that could affect households' ability to share household work (Goldin, 2014).

The differences in gender ideologies refer to the issue that it is not obvious how well questions about gender attitudes capture individual preferences rather than perceived social norms that are not necessarily endogenized by the respondent. My results hint that the problem may be that society in post-socialist countries approves of women performing "double roles" (e.g., in Table 2.3). This is in line with prior research that suggests that women from East Germany continue to have more desire to combine work and family than women from West Germany, even though they are more difficult to combine in the market economy (Adler, 2002; Campa and Serafinelli, 2019). Summary statistics and time and policy regime fixed effects from regression analysis also indicate that while gender convergence in time allocation to household work slowly progresses in Western countries, post-socialist countries are lagging behind.

A narrative that could explain this pattern is that women under socialism, despite all the challenges, managed to keep up with their duties at work and at home. To some extent, this required involvement of men in household work to a larger degree than was common in non-socialist countries at that time. As the decades passed, and Western countries experienced convergence in both gender labor force participation rates and allocation of household work, little has changed in socialist countries. When the transition started, the allocation of unpaid work at home may have been more equal in post-socialist countries than in many advanced economies. However, as gender convergence in the West continues, the number of countries with unpaid work-at-home more equally allocated than in post-socialist countries will grow. If this is the case, the emancipation inspired (or forced) by the socialist regimes may see its reversal during the post-socialist period.

Chapter 3

Inflation Expectations and Labor Supply: Evidence From an Experimental Study

Joint with ChaeWon Baek

> *"Inflation has just about everyone's attention right now, which highlights a particular risk today: The longer the current bout of high inflation continues, the greater the chance that expectations of higher inflation will become entrenched. ... History shows that the employment costs of bringing down inflation are likely to increase with delay, as high inflation becomes more entrenched in wage and price setting."*
>
> — Jerome Powell, at the *Jackson Hole Symposium* on August 26th, 2022.

3.1 Introduction

How do workers change their labor supply decisions in response to changes in expected inflation? The answer to this question is important in order to understand whether and to what extent changes in expected inflation play a role in explaining fluctuations in labor supply over the business cycle. This question is particularly relevant today when many countries experience elevated inflation rates despite the central bank's efforts to curb inflation. In June 2022, the U.S. inflation rate hit its highest level since 1982 of 9.1%. According to the Federal Reserve Bank of New York's Survey of Consumer Expectations, inflation expectations were also running high at 6.8% in June 2022. If wages are, in turn, responsive to the revision of inflation expectations, resulting wage increases could cause prices to rise further. Such dynamics can launch a wage-price spiral, thus making it very difficult for a central bank to control inflation.

Although wage-price inflation is much discussed, as can be seen from the quote above, there has been no direct causal evidence of the relationship between (expected) inflation and labor supply. To test this relationship empirically, one needs information on subjective expectations about the economy and labor supply preferences. Furthermore, while such information can be obtained from observational data, variations in subjective expectations about future economic variables are unlikely to be exogenous. In a similar spirit, individuals' observed labor supply decisions could reflect many unobserved factors researchers cannot directly control for.

We overcome these problems by designing and running an experiment in an online labor market Amazon Mechanical Turk (MTurk, hereafter) in April-July 2022.[1] Specifically, we hire workers to perform a series of forecasting tasks during which we vary their expectations via *randomized* information provision. This allows us to generate exogenous variation in subjective expectations about the economy and thereby identify causal effects on worker's behavior (see Haaland et al., 2023). Specifically, we examine how the resulting revision of expectations affects labor supply measured by reservation wages and desired employment duration. The main advantage of conducting the experiment in MTurk is that, in addition to asking *hypothetical* questions about labor supply, we can credibly offer workers employment on the terms provided by respondents by following up with them based on their answers. Therefore, we can capture the *actual* labor supply response in the online labor market.

The experiment shows that information treatments affect participants' expectations about price inflation, wage inflation, and unemployment rates. Participants meaningfully updated their subjective forecasts based on the provided information. When respondents received one relevant signal, they updated their expectations about all variables jointly. For example, respondents updated their wage inflation expectations and unemployment expectations when provided with the current CPI inflation rate. Similarly, they updated their price inflation and unemployment rate expectations when they received information about hourly earnings inflation rates. This suggests that researchers need to control for all observed expectations *jointly* to avoid potential omitted-variable biases when examining the effect of expectations on households' behaviors. We also find that information treatments affect beliefs across all three waves.

The variation in expectations due to randomized information treatment allows us to analyze the causal relationship between inflation expectations and labor supply in a cross-section of respondents. We elicit labor supply preferences by asking about desired pay and duration of employment for working on a similar task with us. We find that in response to exogenous variation in macroeconomic expectations, MTurk workers adjust their labor supply preferences. Specifically, when they update their hourly earnings inflation expectations upwards, they increase their reservation wages. Moreover, they appear to be more willing to switch to other employers (MTurk requesters) and increase the number of hours worked in their day jobs. In contrast, higher unemployment expectations significantly increase the desired duration of employment with us but do not change reservation wages. Lastly, when workers adjust their price inflation expectations upwards, they rather *decrease* their reservation wages but do not change their desired du-

[1]We received the IRB approval from Tufts University (STUDY00002463) and University of California at Berkeley (IRB 2022-01-14981).

ration of employment. We associate this decrease in reservation wages with the stagflationary view of U.S. households from our first stage results about the information treatment effects. When provided with the current CPI inflation rate, which tends to be higher than their expectation, respondents further increase their unemployment expectations. That is, households associate higher inflation with a bad economic outlook consistent with Kamdar (2018) and Binder (2020). This induces them to reduce the smallest reward necessary for accepting a job offer.

Overall, our results suggest that, contrary to policymakers' concerns, the risk of the wage-price spiral in the U.S. is limited. Even though current high inflation could raise price and wage inflation expectations, this would likely increase unemployment expectations at the same time. While higher wage inflation expectation raises reservation wages, higher price inflation and unemployment expectations decrease reservation wages at the same time, partially offsetting the initial shock. Moreover, in the face of higher wage inflation expectations, workers tend to increase the labor hours of their day jobs. This suggests that wage-price spirals do not seem to be very likely.

To our knowledge, this is the first study to empirically examine the *direct* causal relationship between inflation expectations and labor supply. More broadly, our paper contributes to a growing literature about the information effect of macroeconomic expectations (see, for example, Coibion et al., 2022b, 2023; Binder, 2020; Cavallo et al., 2017; Coibion et al., 2021, 2022a; Hajdini et al., 2022a) and the effects of macroeconomic expectations on behavior (see, for example, Armona et al., 2019; Armantier et al., 2016; Bottan and Perez-Truglia, 2020; Coibion et al., 2022b, 2023; Hajdini et al., 2022b; Belot et al., 2022). These studies have shown that randomized information treatment can successfully generate exogenous variation in households' inflation expectations. A distinguishing feature of our experiments compared to these studies is that we implemented them in a high inflation period when workers have more incentives to be informed about inflation (even so, we find clear treatment effects of information provision on inflation expectations). By building on this rapidly growing literature, we provide novel evidence on the effect of expected inflation on labor supply decisions.

We also contribute to the literature studying wage-price inflation spirals and the role of expectations in generating these spirals. In short, labor market developments depend on how workers form their expectations and adjust their labor supply accordingly. Previous empirical studies have relied on observational data across different countries (see, for example, Kandil, 2003; Boissay et al., 2022). However, because of the inherent endogeneity of subjective expectations, the available evidence is not identified cleanly. We use RCT to generate exogenous variation in subjective expectations and hence our results provide direct causal evidence.

Clearly, understanding how households adjust their labor supply to expected inflation is important for policy discussions and communications. For example, many central banks have made enormous efforts to control inflation expectations. Our results could provide evidence of the direct effects such policies might have on labor supply. Our study is particularly relevant today. With elevated inflationary pressures, workers are more likely to pay attention to changes in inflation and adjust their behavior accordingly. In this regard, our results could help design employment policy by providing useful guidance on likely changes in labor supply in this high-inflation environment.

The remainder of the paper is organized as follows. Section 3.2 describes survey and experimental design. Section 3.3 presents the treatment effects of information provision on subjective expectations. Section 3.4 then examines how changes in expectations affect labor supply preferences. Section 3.5 discusses robustness of results to alternative specifications. Lastly, section 3.6 concludes.

3.2 Survey and Experimental Design

This section describes the survey and experimental design we use to elicit the effect of inflation expectations on labor supply and provides descriptive statistics of participants. Our study design follows recommendations in Haaland et al. (2023).

3.2.1 Survey Design

We implemented our survey via Amazon Mechanical Turk (MTurk). Amazon MTurk is a crowdsourcing website for hiring remotely-located crowd workers to perform on-demand tasks, called HITs (Human Intelligence Tasks), in exchange for monetary rewards. We posted our HITs on MTurk in April and May 2022 for the first wave of our survey. We informed participants that the purpose of the HIT was training a machine learning forecasting algorithm in order to motivate them to carefully answer forecasting questions and avoid the experimenter demand effects. For the quality of data, we allowed participation only for those age 18 or older who had completed at least 1,000 HITS on MTurk and had approval rates of at least 75%.[2] Because our information treatment is for the U.S. economic variables, we restrict our sample to residents of the U.S. (*i.e.* those registered at MTurk in the U.S. and having a U.S. location.) No additional demographic criteria were applied for sample selection. A total of 10,758 MTurk workers (MTurkers, hereafter) attempted to participate in our survey. Among them, 5,487 MTurkers completed the first wave of our survey.[3]

Our survey consists of six blocks. Figure 3.1 summarizes our survey flow. The survey begins with a screening task and a numerical competence check. They are followed by the main part of the survey which allows us to compare the initial forecasts and labor supply preferences with their revised version. The revision of expectations and labor supply preferences is prompted by the randomized information provision in the "Main task".

[2] Requesters who post HITs approve MTurkers' HIT submissions based on their answers. If their answers meet certain criteria set by each requester, they approve HITs. Once their HITs are approved, MTurkers receive posted rewards. Otherwise, they will not receive any rewards.

[3] Attrition from the attempt to the completion is not systemically correlated with the treatment arms (see Appendix Table C.1 and C.2 for details).

In the "Main Task", a key element of our experimental design, we provide random subgroups of respondents with different information about price and wage inflation rates and unemployment which allows us to generate exogenous variation in expectations and thereby to identify the causal effect of expectations revision on labor supply. At the end of the survey, respondents are asked to provide some basic demographic information as well as additional information about their employment offline and online. The specific questions asked are available in Appendix C.7.

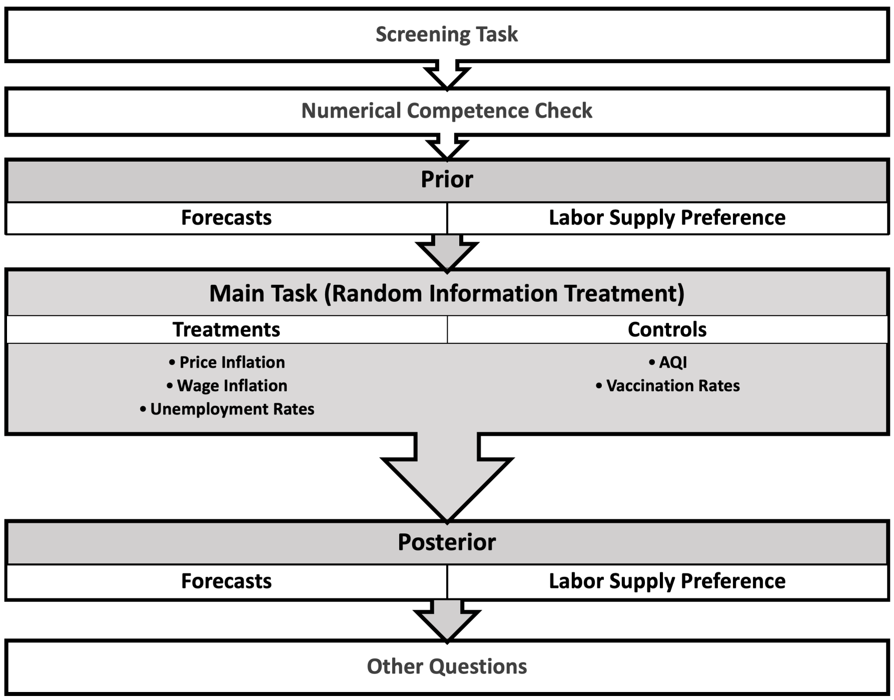


Figure 3.1: Survey flow

Screening Task. Our survey starts with a screening task. The screening task is of a similar format to the main task related to the information treatment. It tests participants' ability to transcribe information from a screenshot accurately. If participants answered the screening task incorrectly, they are prompted to the end of the survey. If the answer is correct, they are prompted to participate in the rest of the survey. We include the screening task to make sure that only those who thoughtfully provide their best answers participate in our survey. Among 10,758 MTurkers who attempted to participate in our survey, 7,457 of them passed the screening task. Among them, 5,487 completed the first wave of the survey. Because most of the attrition happened early in the survey, due to inaccurate answers to screening tasks or reluctance to complete numerical competence checks, attrition is not systemically correlated with the information treatment.

Numerical Competence Check. Upon successful completion of the screening task, participants are prompted to solve a few mathematical problems that evaluate their numerical competence. These questions are designed to check respondents' ability to convert pay per 10 minutes to hourly pay and evaluate percentage change based on absolute change. Although respondents answered these questions incorrectly, they were still able to proceed and complete our survey. Because we provided the information treatments (price and hourly wage) in change *rates* and pay respondents per 10 minutes of work, we include these questions to learn how many respondents are comfortable with interpreting such information. In our sample, about 87% of the participants answered at least two questions correctly. About 75% of the participants answered all three questions correctly.

Prior. This block consists of questions about forecasts and labor supply preferences. Before providing participants with any additional information about macroeconomic variables, we asked for their subjective forecasts for the following variables: price inflation rates, hourly earnings inflation rates, unemployment rates, air quality index in Seattle, and COVID-19 vaccination rates. These variables are associated with our randomized information treatment. In addition to this, we elicited on what terms (desired duration and reservation rewards) respondents were willing to accept and complete follow-up HITs. First, we asked what was the smallest reward for a respondent to be willing to accept a similar HIT taking *10 minutes* of their time *per month* using the following question:

> "*Suppose after completing a HIT on MTurk you are offered to participate in a follow-up task that asks you to do a 10-minute HIT two times – in May and June 2022. What is the smallest reward for 20 minutes of your work that you would accept? (in USD)*"

We then asked for how many months a respondent would be interested in accepting a similar HIT using the following question:

> "*Suppose you could choose for how many months to work on a monthly hit paying (*a respondent's own answer for the reservation wage question*) USD for 10 minutes of work. For how many months would you prefer to work?*"

Main Task. In this block, we randomly assign MTurkers into one of the five groups: three treatment groups and two control groups. Each group is provided with different information treatment in the form of a text transcription task. Specifically, respondents are asked to transcribe information from the screenshot into a table. The information refers to official information about either macroeconomic variables of interest (price inflation, hourly earnings inflation, and unemployment rate – treatment groups) or variables unrelated to a macroeconomic situation (air quality index in Seattle and Covid-19 vaccination rates – control groups). Our identification strategy exploits exogenous variation in macroeconomic expectations for respondents in the treatment groups, i.e., provided with pertinent information, relative to those in the control groups. The examples of screenshots are available in Appendix C.6. For instance, participants assigned to a price inflation group were prompted to a screenshot of the BLS report about Consumer Price Index (CPI) inflation (Appendix Figure C.3). They were asked to transcribe the data about the CPI 1-month percentage change and 12-month percentage change. Similarly, participants assigned to a wage inflation group were prompted to transcribe the average hourly earnings in the private sector in the U.S. from a BLS news release (see Appendix Figure C.4). To ensure that participants paid attention to the information treatment, they were informed that if they recorded the information from the screenshot incorrectly, they would not be paid for the entire HIT. We also added attention-check questions to verify the recall rate after completion of the transcription task. About 75% of the participants in the price and wage inflation treatment groups correctly recalled the information they transcribed.

Posterior. After the information treatment, we elicited respondents' subjective expectations about the economy (price and hourly earnings inflation rates and unemployment) and other variables in the control group (air quality in Seattle and Covid-19 vaccination rates) again. We used similar but different wording to avoid asking exactly the same questions. We then asked about their desired duration of employment and reservation wages again. Specifically, we used the following questions similar to those in the prior block:

> "*Suppose in the future we offered you to perform a* similar *task you did today taking about 10 minutes of your time once a month. I.e. you would record the information from the same website and provide your prediction based on it. How many months would you be interested in working?*"
>
> "*In the previous question, you answered that you are willing to work on a similar 10-min task for (*a respondent's own answer to the previous question*) months, which corresponds to (*10×a respondent's own answer to the previous question*) min of your time. What is the **lowest** total reward that you would accept to work? (in USD)*"

Other Questions. In this block, we asked about respondents' characteristics such as gender, age, education level, employment status, household income, marital status, number of children, etc. Furthermore, we asked some hypothetical labor supply questions for their day jobs in *offline* labor markets. Answers to these questions complement our main analysis of labor supply preferences in the online labor market.

3.2.2 Follow-up Surveys

At the beginning of the survey, respondents were informed that our HIT is designed to train a machine-learning algorithm for forecasting. This description signals to participants that answers to forecasting questions are very important for the project's success, but it is different than the "true" purpose of the survey, which is to examine how the revision of people's subjective expectations affects their labor supply decisions. We chose not to fully disclose the purpose of our study for the following reasons. First, the full disclosure of the survey's purpose could bias respondents' responses about labor supply decisions. Second, we wanted MTurkers to understand that our project is an ongoing project that takes a few months with follow-up HITs. Because MTurk is an *actual* labor market, we expected them to believe that we would follow up with them based on their answers for the desired terms (rewards and duration), thereby providing us with their best answers. This would allow us to learn about their labor supply preferences without asking *hypothetical* questions.

Based on their answers in the first wave, we followed up with respondents interested in participating in the follow-up HITs. If participants answered that they would be willing to participate in the follow-up HITs, we offered them an opportunity to work with us in the following month at the rate they asked for. Among 3,979 participants in wave 1, net of duplicates, we followed up with 2,763 participants: those in the two treatment groups (CPI and hourly earnings group) and those in the AQI control group. Among them, about 1,450 (about 52%) participated in the second and/or third waves, and 937 of them participated in all three waves.[4]

3.2.3 Descriptive Statistics

Table 3.1 provides descriptive statistics about respondents. In terms of gender, race, and age, our sample is representative of the U.S. population. The average age is about 40 years old, about half of them are female, and 80% of them are white. But our respondents are more educated compared to the U.S. population, as other MTurkers are.[5] About 75% of them have a 4-year college degree or more. About 83% of them are either employed full-time or employed part-time. In other words, most of them have day jobs and not many of them use MTurk as their major income source. Nonetheless, they spend on average 20.39 hours per week working on MTurk. Their households spend \$724 for food and \$290 for gas per week. The median household income bin is \$50,000 - 59,999 per year.

The average expected price inflation rate is 6.2% and the median expected inflation rate is 5%. According to the Michigan survey of consumer sentiments, the median one-year ahead inflation expectation was 5.4% in April 2022 and 5.3% in May 2022. The median expected inflation rate from the New York Fed's survey of consumer expectations

[4]Appendix Table C.3 summarizes attrition from participation in the follow-up waves of the survey.

[5]Our survey has numerical competency check questions. It is more likely that those who are more comfortable with numbers tend to complete our surveys.

is 6.3% in April and 6.6% in May. The average and median from our survey are close to these numbers but are lower than the actual inflation rate of around 8% in April and May 2022. The average expected wage inflation rate is 7.20%, which is higher than the actual wage inflation rate of around 5% in April and May 2022. But the median expected wage inflation rate is 4%, which is lower than the actual wage inflation rate. The average expected unemployment rate is 7.2% which is more than double the actual unemployment rate of around 3.5% in April and May 2022.[6] The average desired duration of employment on a monthly HIT like ours is 3.78 months, and the average reservation wage is about $1 per 10 minutes of work. Descriptive statistics about respondents in the second and the third waves are similar to Table 3.1 (see Appendix Table C.4)

Table 3.1: Descriptive statistics (late April-May, 2022)

	Mean	Percentiles			Std. Dev.
		p25	p50	p75	
age	40.35	31.00	39.00	48.00	11.98
female	0.50	0.00	0.00	1.00	0.50
white	0.80	0.00	1.00	1.00	0.40
with college degree	0.75	0.00	1.00	1.00	0.43
employed	0.83	0.00	1.00	1.00	0.38
full-time employed	0.69	0.00	1.00	1.00	0.46
number of children	0.92	0.00	1.00	2.00	1.07
monthly spending on food	$723.91	$150.00	$350.00	$600.00	$2742.63
monthly spending on gas	$289.96	$40.00	$100.00	$200.00	$1774.25
$\mathbb{E}_t^{\texttt{prior}}[\pi_{t+12}]$	6.16	1.00	5.00	10.00	8.18
$\mathbb{E}_t^{\texttt{prior}}[\pi^w_{t+12}]$	7.20	1.00	4.00	10.00	11.27
$\mathbb{E}_t^{\texttt{prior}}[u_{t+12}]$	7.20	4.45	6.45	9.13	3.75
$\Delta^{\texttt{post-prior}}\mathbb{E}_t[\pi_{t+12}]$	0.53	-2.00	0.00	3.00	7.56
$\Delta^{\texttt{post-prior}}\mathbb{E}_t[\pi^w_{t+12}]$	-1.00	-3.00	0.00	2.00	11.60
$\Delta^{\texttt{post-prior}}\mathbb{E}_t[u_{t+12}]$	0.84	-1.16	0	1.83	4.91
$\mathbb{E}_t^{\texttt{prior}}[\texttt{duration}_{t+1}]$	3.78	3.00	5.00	5.00	1.53
$\mathbb{E}_t^{\texttt{prior}}[\texttt{reservation wages}_{t+1}]$	0.99	0.50	1.00	1.25	0.54
Observations	3,979				

[6]When we asked about their expected unemployment rates, we gave information about the lowest and highest unemployment rates between 2019 and 2021.

3.3 Effects of Information Provision on Subjective Expectations

This section studies the treatment effect of the information provision. Before and after the information treatment, respondents were asked about their subjective expectations about macroeconomic and other variables. Based on this information, we study if respondents update their expectations when they receive a relevant signal relative to an irrelevant one. We are interested in whether there are systematic differences in the revision of expectations across treatment groups relative to the control groups. Since respondents were randomly allocated into treatment vs control groups, the differential revision patterns must be caused by the information signal they received. To illustrate the expectations revision, we first analyze binned scatter plots of respondents' posterior price inflation expectations and their revisions against their priors, and then perform regression analysis.

3.3.1 Graphical Representation

Each panel of Figure 3.2 summarizes the revision of expectations about one macroeconomic variable due to information treatment. Panel A presents the results for price inflation expectations for respondents in the treatment group who received information about the CPI inflation rate and respondents in the control group who received information about the air quality index and Covid-19 vaccination rate. If respondents in the treatment group did not pay attention to the information about inflation they received, they should behave in a similar way as the control group that received information largely irrelevant to macroeconomic conditions. The revision of price inflation expectations in the control group can be attributed to a change in wording in prior and posterior questions. The difference between revision patterns in the control and treatment groups illustrated by black and blue lines respectively illustrate the effect of the information treatment.

The left graph of panel A shows that respondents who have received the relevant information about the current CPI inflation rate exhibit a much flatter slope compared to those in the control group who have received irrelevant information. This suggests that, in line with Bayesian updating, those in the treatment group place much smaller weights on their priors.[7] The right graph of panel A points to a similar conclusion.

[7]To illustrate belief updating, consider a worker with a prior expectation of macroeconomic variable of interest $\mathbb{E}^{\text{prior}}[Z_{t+12}]$ who receives a relevant *Signal*. Under Bayesian learning, workers' posterior expectation should be a weighted average of a prior and a signal:

$$\mathbb{E}^{\text{post}}[Z_{t+12}] = (1-\alpha)\mathbb{E}^{\text{prior}}[Z_{t+12}] + \alpha \; Signal$$

and revision of expectations should be a similar function of a prior and a signal:

$$\mathbb{E}^{\text{post-prior}}[Z_{t+12}] = \alpha \; Signal - \alpha \; \mathbb{E}^{\text{prior}}[Z_{t+12}]$$

The graphical and regression specifications in text estimate weight parameter α from the equations above.

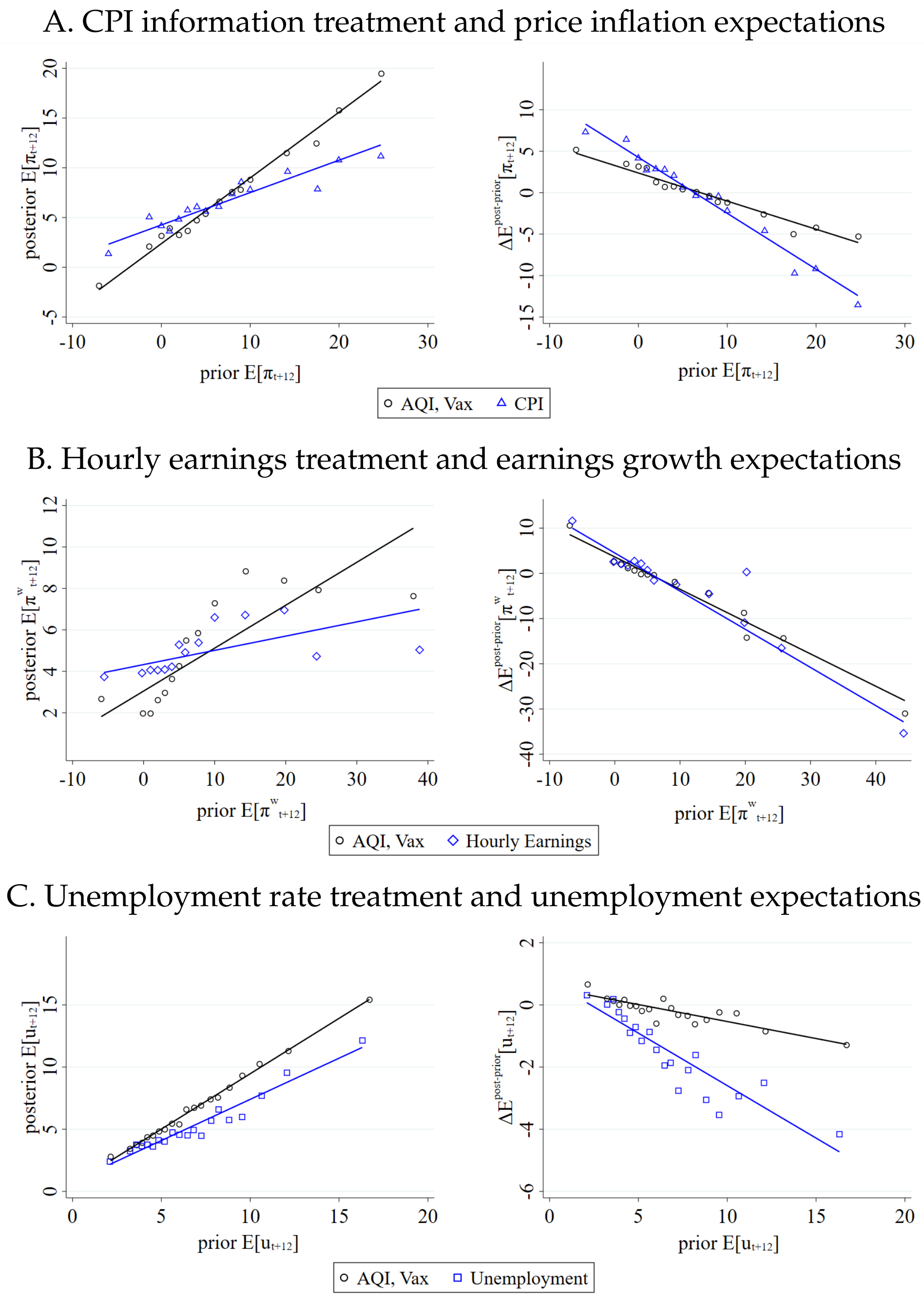


Figure 3.2: Effects of information treatments on macroeconomic expectations

Notes: This figure draws binned scatter plots of the highly numerate respondents' posterior expectations over the next 12 months (the left panel, on *y*-axis) and their revision of forecasts (the right panel, on *y*-axis) against their priors (on $x-$axis) to illustrate the effect of the most relevant information provision from the first wave of the survey. Huber-robust weights are applied. Highly numerate respondents are those who answered all numerical competence check questions correctly. Additional results for revision of expectations in response to various signals are reported in Appendix C.2.1.

As depicted by the difference in slopes of black and blue lines, respondents in the control group revise their expectations less than those in the treatment group given their prior inflation expectations. Taking into account that those in the treatment group were provided with a signal about an annual CPI inflation rate of 7.9%, the graph shows that respondents revise their expectations toward the signal by placing a higher weight on the signal and decreasing weight on the prior. Panels B and C of Figure 3.2 depict a revision of earnings growth expectations and unemployment expectations in response to signals about past earnings growth and unemployment rate forecast. Similar to the results in Panel A, respondents who received relevant information placed a lower weight on their priors and a higher weight on the signal than those in the control group.

Since many macroeconomic phenomena are interrelated, revisions of macroeconomic expectations about one variable may be responsive to signals about other macroeconomic variables. To examine whether this is the cause in our experiment, we analyze how respondents revise their expectations about a variable Z (e.g., price inflation expectations) when they receive a signal about another variable (e.g., wage growth rate or unemployment rate). Appendix Figure C.1 indicates that a signal about hourly earnings growth results in a similar revision of price inflation expectations as a signal about the CPI inflation rate. The effect of a signal about the unemployment rate is qualitatively similar, although smaller in magnitude. Similar to price inflation expectations, hourly earnings growth expectations react to signals about several variables (Appendix Figure C.2). At the same time, unemployment rate expectations are largely unresponsive to signals about price and wage inflation.

3.3.2 Regression Analysis

To study the effect of information treatments on expectations revision more formally, we analyze the effect of information treatments illustrated in Figure 3.2 by estimating the following regression equation:

$$\begin{aligned}\Delta\mathbb{E}_{it}^{\text{post}-\text{prior}}[Z_{t+12}] = \alpha_0 + \alpha_1\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}] + \alpha_2\texttt{treat}_i^Z + \\ + \alpha_3\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}] \times \texttt{treat}_i^Z + \varepsilon_i \qquad (3.1)\end{aligned}$$

for $Z = \{\pi, \pi^w, u\}$. Here, $\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}]$ is a prior expectation of variable z over the next 12 months, $\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$ is a revision in expectations for variable Z after the information provision, and $\texttt{treat}_i^Z$ is a treatment dummy denoting if a respondent i is in the treatment group that received a signal about variable Z. In other words, to study information treatment effects, we regress forecast revisions following the information treatment on prior expectations, treatment dummy, the interaction between a treatment dummy and prior expectation, and a set of control variables. Following Coibion et al. (2023), Coibion et al. (2022b), Hajdini et al. (2022b) and others, we use Huber-Robust regressions to control for outliers. The results are summarized in Table 3.2.

Columns 1-3 of Table 3.2 show the effect of information treatment about the CPI inflation rate on the revision of inflation expectations. First, when provided with information about the current CPI inflation rates, respondents, on average, revise their posterior expected price inflation rates upward by 1.5-1.9 percentage points. In addition, their implied weight on prior price inflation expectations falls from 0.66-0.70 by 0.30-0.34 points. The results in columns 4-6 show that statistics about hourly earnings have a statistically significant effect on wage inflation expectations. Respondents, on average, revise their posterior expectations by 1.1-2.2 percentage points as well as reduce weight on prior wage inflation expectations from 0.19-0.45 by 0.13-0.40 points. Finally, according to columns 7-9, when workers received information about the unemployment rate forecast, they updated their unemployment rate expectations upward by 0.22-0.96 percentage points. In addition, they reduced the weight they put on prior from 0.89-0.91 by about 0.23-0.36 points. These results support the conclusion that information treatments induce respondents to revise their expectations as intended.

Table 3.2: Effects of information treatments on the revision of price inflation, wage inflation, and unemployment expectations

Dependent variable:	Price inflation ($Z = \pi$)			Wage inflation ($Z = \pi^w$)			Unemployment rate ($Z = u$)		
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
`treat_cpi`	1.88***	1.54***	1.45***						
	(0.25)	(0.26)	(0.27)						
`treat_wage`				1.28***	1.13***	2.17***			
				(0.20)	(0.21)	(0.22)			
`treat_unemp`							0.22	0.40*	0.96***
							(0.23)	(0.24)	(0.24)
`treat_cpi` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.34***	-0.31***	-0.30***						
	(0.02)	(0.02)	(0.03)						
`treat_wage` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$				-0.14***	-0.13***	-0.37***			
				(0.02)	(0.02)	(0.02)			
`treat_unemp` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$							-0.23***	-0.25***	-0.36***
							(0.03)	(0.03)	(0.03)
$\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.34***	-0.31***	-0.30***	-0.79***	-0.81***	-0.55***	-0.11***	-0.10***	-0.09***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	2406	2352	1766	2468	2409	1777	2093	2042	1533
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output from equation (3.2). For each outcome variable specified in the header, the first column reports results without controls, the second column adds control variables, and the third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates.

Specification (3.1) implies that respondents revise expectations about a specific variable only if they receive a signal about this variable. However, given one signal, respondents may revise multiple expectations simultaneously (see Appendix C.2.1). If this is the case, results in Table 3.2 suffer from an omitted variable bias. To avoid the bias, and allow for the possibility that multiple information treatments affect expectations for multiple variables, we extend equation (3.1) by including indicator variables for multiple information treatments and their interactions with the prior expectation of the variable of interest for $Z = \{\pi, \pi^w, u\}$.

$$\begin{aligned}\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}] =& \beta_0 + \beta_1\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}] \quad + \sum_{k\in\{\pi,\pi^w,u\}} \beta_{2,k}\texttt{treat}_i^k \\ &+ \sum_{k\in\{\pi,\pi^w,u\}} \beta_{3,k}\left(\texttt{treat}_i^k \times \mathbb{E}_{it}^{\text{prior}}[Z_{t+12}]\right) + \mathbf{X}_i'\gamma + \varepsilon_i \end{aligned} \tag{3.2}$$

The estimation results for equation (3.2) are reported in Table 3.3. Columns 1-3 show the effect of information treatments on the revision of inflation expectations. When provided with information about the current CPI inflation rates, respondents revise their posterior expected price inflation rates upward by 1.5-1.9 percentage points, about the same amount as in Table 3.2. When respondents receive information about the current hourly earnings inflation rate, they also increase their expected *price* inflation rate on average. Similarly, respondents in the treatment group place significantly smaller weights on their priors than those in the control group, both when provided information about the CPI inflation rate and about other macroeconomic variables (the negative and statistically significant coefficient on $\texttt{treat} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}]$). In other words, respondents in the treatment group update their expectations when provided with *any* relevant signal. This is consistent with earlier works on the effects of information treatment on inflation expectations (see, for example, Coibion et al., 2023, 2022b; Binder, 2020; Cavallo et al., 2017; Hajdini et al., 2022b). It also shows that respondents update their subjective expectations about future price inflation not only in response to the signal about current inflation rates but also in response to other relevant information such as hourly earnings inflation rates and unemployment rates. But they are more responsive to the direct signals about the current CPI inflation rate and/or hourly earnings wage inflation, rather than to signals about unemployment rates.

We observe similar patterns for hourly earnings inflation expectations from columns 4-6 of Table 3.3. First, respondents, on average, increase their expected wage inflation rates when they are provided with either the current CPI inflation rate or the hourly earnings inflation rate. When given the information about the current hourly earnings inflation rate, they increase their expectations about hourly earnings inflation rates by 1-2 percentage points. When they are provided with information about the current CPI inflation rates, they increase their expected hourly earnings inflation rates by 0.6-1.2 percentage points. Second, respondents in the treatment group place significantly smaller weights on their priors than respondents in the control group. The implied weight on the prior expectations falls from 0.20-0.45 by 0.14-0.31. Similar to price inflation expectations, respondents react not only to the most relevant information, signal about hourly earnings inflation, but also to CPI inflation rates and unemployment rates. This suggests that when receiving information about the current CPI inflation rates, respondents update not only their expectations about price inflation rates but also other macroeconomic expectations. As is the case for the CPI inflation expectations, however, they are more responsive to the signal about price or hourly earnings inflation than the signal about unemployment rates.

Table 3.3: Effects of information treatments on the revision of price inflation, wage inflation, and unemployment expectations (multiple treatments)

Dependent variable:	Price inflation ($Z = \pi$)			Wage inflation ($Z = \pi^w$)			Unemployment rate ($Z = u$)		
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
`treat_cpi`	1.90***	1.56***	1.48***	0.78***	0.60***	1.21***	-0.17	-0.21	-0.23
	(0.25)	(0.26)	(0.27)	(0.22)	(0.22)	(0.25)	(0.23)	(0.24)	(0.22)
`treat_wage`	1.47***	1.23***	1.47***	1.27***	1.09***	1.93***	-0.15	-0.15	-0.14
	(0.23)	(0.24)	(0.26)	(0.21)	(0.21)	(0.23)	(0.23)	(0.24)	(0.24)
`treat_unemp`	0.02	0.04	-0.07	-0.46*	-0.40	0.33	-0.04	0.00	0.74***
	(0.28)	(0.29)	(0.30)	(0.24)	(0.25)	(0.27)	(0.27)	(0.28)	(0.26)
`treat_cpi` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[z_{t+12}]$	-0.34***	-0.31***	-0.30***	-0.13***	-0.13***	-0.30***	0.11***	0.10***	0.08***
	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
`treat_wage` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.30***	-0.29***	-0.32***	-0.14***	-0.14***	-0.31***	0.05*	0.05	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
`treat_unemp` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.14***	-0.13***	-0.09***	-0.03*	-0.05**	-0.21***	-0.18***	-0.18***	-0.31***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
$\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.34***	-0.35***	-0.27***	-0.79***	-0.80***	-0.61***	-0.12***	-0.11***	-0.09***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	3976	3892	2889	3979	3895	2890	3979	3895	2890
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output from equation (3.2). For each outcome variable specified in the header, the first column reports results without controls, the second column adds control variables, and the third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates.

In contrast to previous results, columns 7-9 of Table 3.3 show that respondents' expectations about unemployment rates mostly respond to the signal about unemployment rates. When provided the information about the current unemployment rates, respondents significantly revise their unemployment rate expectations toward the signal. The implied weight on the prior expectations falls from 0.88-0.91 by 0.18-0.31. When provided with the signal about high current price inflation rates, respondents further corroborate their prior unemployment rate expectations. Interestingly, the positive coefficient on `treat_cpi` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[u_{t+12}]$ in Table 3.3 shows that people put even *higher* weights on their priors when receiving signals about price inflation rates. This is consistent with a stagflationary view of inflation (see, for example, Kamdar, 2018; Binder, 2020). That is, they tend to think that when inflation rates are higher, unemployment rates tend to increase as well.

While information treatments induce respondents to revise macroeconomic expectations in the short run, these effects persist over a longer horizon (see Table C.6 in Appendix C.3.1). Specifically, we find that when respondents update their expectations, they still place some weight on the relevant information that they received one or two months ago. The implied weights on the information received in the past are, however, smaller than weights on information received contemporaneously from Table 3.3. This is consistent with standard Bayesian learning. As time passes, the information gets more dated, so respondents put less weight on the information that they received a month or two months ago. The fact that respondents in the treatment groups have learned about either the current CPI inflation rates or hourly earnings inflation rates by participating in the first wave of the survey could weaken the information treatment effect from the subsequent follow-up surveys. Although we find statistically significant information treatment effects across all three waves, the magnitude of the effect decreases in the third wave (see Table C.7 in Appendix C.3.2). This is consistent with "learning-through-survey" effects documented by Binder and Kim (2020).

To recap, this section studies the effect of information treatment on subjective inflation and unemployment expectations. We find that, on average, respondents increase their posterior price or wage inflation expectations when they are provided with either the current CPI inflation rate or the hourly earnings inflation rate. Interestingly, they update their posterior price (wage) inflation rate upwards even when they receive information about the current hourly earnings (CPI) inflation rates. Moreover, individuals in the treatment groups place significantly smaller weights on their priors than those in the control group. Price inflation expectations respond to both signals about price and hourly earnings inflation. The same is true for hourly earnings inflation expectations. Unemployment rate treatment has larger effects on hourly earnings inflation expectations than on price inflation expectations. Unemployment expectations respond mostly to the signal about current unemployment rates. When provided with information about current high CPI inflation rates, respondents tend to revise their expectations about unemployment rates in the next 12 months *upwards*.

Overall, our results show that when provided with *one* relevant signal, respondents update their expectations about *all* variables altogether. This suggests that when examining the effect of macroeconomic expectations on households' behaviors, we need to control for *all* observed expectations to avoid potential omitted variable biases. For this reason, when we examine how expectations affect labor supply preferences, we include posterior price, wage inflation, and unemployment expectations at the same time.

3.4 Effects of Subjective Expectations on Labor Supply

In this section, we examine the *causal* relationship between expected inflation and labor supply. As we discussed above, subjective expectations about future economic variables are unlikely exogenous. Many unobserved factors affect both expectations and individuals' labor supply decisions. To overcome these issues, we use an instrumental variable approach. In light of the discussion in Section 3.3, we use information treatments and the interactions of information treatments with priors as instruments to identify exogenous variations in expectations and study the causal link between expectations and labor supply decisions.

As we have examined in Section 3.3, when provided with *one* of the relevant pieces of information about the economy, respondents update their expectations about *all* relevant variables together. For example, when respondents received information about CPI inflation rates, they updated their expectations about price inflation rates, wage inflation, and unemployment rates. For this reason, we estimate the regression model with all the measured expectations (price, wage, and unemployment rates) as endogenous variables in the second-stage equation:

$$\begin{aligned}\Delta Y_{it}^{\texttt{post-prior}} &= \beta_0 + \beta_1 \Delta \mathbb{E}_{it}[\pi_{t+12}] + \beta_2 \Delta \mathbb{E}_{it}[\pi^w_{t+12}] + \beta_3 \Delta \mathbb{E}_{it}[u_{t+12}] \\ &\quad + \gamma_1 \mathbb{E}_{it}^{\text{prior}}[\pi_{t+12}] + \gamma_2 \mathbb{E}_{it}^{\text{prior}}[\pi^w_{t+12}] + \gamma_3 \mathbb{E}_{it}^{\text{prior}}[u_{t+12}] + \mathbf{X}'_{it}\delta + \eta_i \end{aligned} \tag{3.3}$$

where $\Delta Y_{it} = \{\Delta \texttt{dur}_{it}^{\texttt{post}}, \Delta \overline{\texttt{rw}}_{it,t+\texttt{dur}_t}^{\texttt{post}}\}$ are changes in desired duration of employment on our MTurk project (in month) and reservation wage per 10-minute monthly task.

Because of the endogeneity inherent in posterior macroeconomic expectation variables in equation (3.3), we instrument them with information treatment dummies and their interactions with prior expectations. The first stage can be concisely summarized with equation (3.2). To be more specific, our instrument set includes the information treatment dummies, the interaction of prior price inflation expectations with the CPI treatment dummy and hourly earnings treatment dummies, the interaction of prior hourly earnings inflation expectations with the CPI treatment dummy and hourly earnings treatment dummy, and the interaction of prior unemployment expectations with unemployment treatment.[8] The parameters of our interest are β_1-β_3's.

3.4.1 Effects on Desired Duration of Employment on MTurk

This section focuses on the effect of macroeconomic expectations on the desired duration of employment on a specific MTurk project. Table 3.4 shows the regression results from equation (3.3) for revision of desired duration of employment on our MTurk project

[8] In other words, we instrument $\Delta \mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$ for $Z \in \{\pi, \pi^2, u\}$ with the following set of IVs: $\texttt{treat\_cpi}_{it}$, $\texttt{treat\_wage}_{it}$, $\texttt{treat\_unemp}_{it}$, $(\texttt{treat\_cpi}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}])$, $(\texttt{treat\_cpi}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi^w_{t+12}])$, $(\texttt{treat\_wage}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi^w_{t+12}])$, $(\texttt{treat\_wage}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}])$, and $(\texttt{treat\_unemp}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[u_{t+12}])$.

within 5-10 minutes before and after information treatment. Columns 1-2 show that in the sample of all respondents, only unemployment rate expectations have a significant effect on the desired duration of employment. Upward revision of unemployment expectations by 1 percentage point increases the desired duration of employment by 3-5 periods. That is, if MTurk workers have a worse economic outlook, they want to increase their labor supply. This result remains significant in the sample of numerate respondents.

Table 3.4: Effects of expectations on desired duration of employment

	Desired Duration (in months)			
	(1)	(2)	(3)	(4)
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[\pi_{t+12}]$	-0.12	-1.26	0.56	0.22
	(1.27)	(1.33)	(0.90)	(0.89)
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[\pi_{t+12}^{w}]$	-0.86	0.01	-2.10**	-1.70**
	(2.09)	(1.94)	(0.70)	(0.76)
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[u_{t+12}]$	4.48***	3.12**	4.41***	2.72**
	(1.48)	(1.46)	(1.21)	(1.19)
N	3,141	3,079	2,222	2,160
Sample	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}]$	11.87	10.42	22.46	20.85
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}^{w}]$	12.18	10.73	50.03	40.53
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[u_{t+12}]$	31.04	29.86	49.91	52.05

Notes: This table presents the regression output to estimate the effects of expectations on the desired duration of employment on our MTurk HIT according to equation (3.3). We instrument the revisions in expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, and the interaction of unemployment treatment dummies with prior expected unemployment rates. Highly numerate respondents are those who answered all the numerical competence check questions correctly. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for details about the treatment of outliers.

We also find that when respondents revise their expected wage inflation rates further upwards, they tend to decrease the desired months of working for us, controlling for expected price and unemployment expectations. This is the case with highly numerate respondents who answered all the numerical competence check questions (calculating percentage changes) correctly. The fact that MTurk workers want to decrease their desired duration of employment suggests two possibilities. First, they want to decrease the overall labor supply. Second, they want to decrease their labor supply for our HITs without decreasing overall working hours in MTurk. For example, Hajdini et al. (2022b) find that when inflation expectation increases, people want to switch to another employer paying higher salaries. To shed light on which scenario is more plausible, we supplement the evidence about labor supply preferences on MTurk analyzed Table 3.4 with additional evidence about the offline labor market preferences elicited by the respondents at the end of the survey (see Section 3.4.3). The results suggest that the second scenario is more likely, and with higher wage growth expectations respondents tend to switch to other employers (MTurk requesters) potentially offering higher-paid HITs without reducing the overall labor supply.

3.4.2 Effects on MTurk Reservation Wages

This section focuses on the effect of macroeconomic expectations on reservation wages in the online labor market. Table 3.5 reports the effect of revision of macroeconomic expectations on the reservation wages per 10 minutes of respondents' time. The data is obtained from an answer to questions we asked before and after the information treatment about the smallest reward that respondents would be willing to accept to complete a similar task in the future.

Table 3.5: Effects of expectations on reservation wages

	Reservation Wages			
	(1)	(2)	(3)	(4)
$\Delta\mathbb{E}_{it}^{\texttt{post}-\texttt{prior}}[\pi_{t+12}]$	-0.50	-0.32	-1.38**	-1.47**
	(0.77)	(0.77)	(0.66)	(0.67)
$\Delta\mathbb{E}_{it}^{\texttt{post}-\texttt{prior}}[\pi^w_{t+12}]$	2.22**	2.88***	1.20***	0.73*
	(0.97)	(1.05)	(0.44)	(0.43)
$\Delta\mathbb{E}_{it}^{\texttt{post}-\texttt{prior}}[u_{t+12}]$	-1.72**	0.17	0.26	0.93
	(0.82)	(0.83)	(0.70)	(0.71)
N	3,075	3,015	2,110	2,056
Sample	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}]$	10.84	10.51	15.68	15.01
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi^w_{t+12}]$	12.71	11.61	54.64	37.89
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[u_{t+12}]$	36.80	31.37	48.21	45.64

Notes: This table presents the regression output to estimate the effects of expectations on reservation wages in the online labor market according to equation (3.3). We instrument the revisions in expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, and the interaction of unemployment treatment dummies with prior expected unemployment rates. Highly numerate respondents are those who answered all the numerical competence check questions correctly. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for details about the treatment of outliers.

The results in Table 3.5 show that respondents raise reservation wages in response to the increase in expected wage inflation rates, after controlling for expected price inflation rates and expected unemployment rates. A one percentage point increase in the expected wage inflation rate is associated with a 1-3 cent increase in their reservation wages per ten minutes. This corresponds to a 1 to 3 *percent* increase given the average/median reward per 10 minutes of $1. On the other hand, higher expected price inflation rates tend to rather *decrease* reservation wages, controlling for expected wage inflation rates and expected unemployment rates in specifications with highly numerate respondents. A 1 percentage point increase in the expected price inflation rate is associated with a 1.3-1.5 cent *decrease* (1.3-1.5%) in nominal reservation wages on average.

We interpret the qualitatively different responses of workers to wage and price inflation as evidence that households have a stagflationary view of inflation, i.e., they interpret inflation as a bad signal about the economy. Therefore, rather than demanding that employers compensate them for the decline in purchasing power of their earnings, they are willing to accept lower pay to secure employment. Importantly, due to the countervailing effect of inflation expectations on reservation wages, such behavior is unlikely to result in a wage-price spiral.

3.4.3 Effects on Offline Labor Supply

The previous discussion focuses on the effect of macroeconomic expectations on *online* labor supply preferences. This section complements these results by examining the effect on preferences in *offline* labor markets. We elicited offline labor supply preferences by asking additional questions at the end of the survey. For the sake of survey time, we did not ask respondents about offline labor supply before the information treatment, which limits the amount of variation available relative to the previous analysis.

We asked respondents about offline labor supply along both extensive and intensive margins. For the extensive margin, we asked respondents to elicit subjective probabilities of changes in labor market status in the next 4 months (e.g., being employed with the same employer, changing employers, becoming self-employed, becoming unemployed, or exiting the labor force). Table 3.6 reports the results. Not surprisingly, respondents with higher unemployment rate expectations have a significantly lower subjective probability of being employed both in the overall sample and in a subsample of numerate respondents. Respondents with higher wage inflation expectations tend to have higher subjective probabilities of being employed, and the result is statistically significant for numerate respondents. At the same time, respondents with higher price inflation expectations are pessimistic about their chances of being employed, especially the numerate ones. These results are consistent with a story that households interpret an increase in inflation as an indicator of deteriorating economic conditions.

Table 3.6: Effects of macroeconomic expectations on probability of employment

	Prob. of Employed			
	(1)	(2)	(3)	(4)
$\mathbb{E}^{\text{post}}_{it}[\pi_{t+12}]$	0.23	-0.29	-0.51	-0.52*
	(0.45)	(0.26)	(0.37)	(0.30)
$\mathbb{E}^{\text{post}}_{it}[\pi^{w}_{t+12}]$	0.35	0.22	0.56***	0.29*
	(0.57)	(0.17)	(0.21)	(0.17)
$\mathbb{E}^{\text{post}}_{it}[u_{t+12}]$	-5.95***	-1.57***	-4.18***	-1.79***
	(0.69)	(0.35)	(0.45)	(0.38)
N	2,870	1,944	1,985	1,938
Sample	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes
F-stat for $\mathbb{E}^{\text{post}}_{it}[\pi_{t+12}]$	12.02	17.27	15.44	14.88
F-stat for $\mathbb{E}^{\text{post}}_{it}[\pi^{w}_{t+12}]$	10.27	40.81	35.21	35.86
F-stat for $\mathbb{E}^{\text{post}}_{it}[u_{t+12}]$	18.71	25.80	32.75	27.09

Notes: This table presents the regression results for the effect of macroeconomic expectations on the subjective probability of being employed or self-employed in the next 4 months according to the following equation:

$$\begin{aligned} P_{it}(\texttt{employed}) = &\beta_0 + \beta_1\mathbb{E}^{\text{post}}_{it}[\pi_{t+12}] + \beta_2\mathbb{E}^{\text{post}}_{it}[\pi^{w}_{t+12}] + \beta_3\mathbb{E}^{\text{post}}_{it}[u_{t+12}] \\ &+ \gamma_1\mathbb{E}^{\text{prior}}_{it}[\pi_{t+12}] + \gamma_2\mathbb{E}^{\text{prior}}_{it}[\pi^{w}_{t+12}] + \gamma_3\mathbb{E}^{\text{prior}}_{it}[u_{t+12}] + \mathbf{X}'_{it}\delta + \varepsilon_i \end{aligned}$$

We instrument the posterior expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, and the interaction of unemployment treatment dummies with prior expected unemployment rates. Highly numerate respondents in columns 3-4 are those who answered all the numerical competence check questions correctly. Heteroskedasticity-robust-standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for details about the treatment of outliers.

Offline labor supply preferences along the intensive margin refer to a desire to change the number of hours worked per week. It is obtained from questions about how many hours respondents work per week on day jobs, whether they would like to change those hours, and by how much. According to Table 3.7, workers with higher unemployment expectations are more likely to be interested in increasing hours worked per week, likely due to precautionary mechanisms. A one percentage point increase in expected unemployment rates increases the probability to desire more working hours by 3-5 percent. A one percentage point increase in inflation expectations has no effect or decreases the probability to desire more working hours by up to 2 percent. As expected, as wage inflation increases, respondents want to work more in their day jobs. A one percentage point increase in wage inflation expectations increases the probability of desiring more working hours by one percent. This adjustment is likely driven by an interplay of income and substitution effects. Combined with the result in Table 3.4 that respondents decrease the desired duration of employment with us due to higher wage growth expectations, the response of offline labor supply preferences suggests that higher wage growth expectations appear to cause *reallocation* of labor supply across employers rather than an overall decrease in labor supply.

Table 3.7: Effects of macroeconomic expectations on desired hours worked

	Increase Hours			
	(1)	(2)	(3)	(4)
$\mathbb{E}_{it}^{\text{post}}[\pi_{t+12}]$	-0.00	-0.01	-0.00	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)
$\mathbb{E}_{it}^{\text{post}}[\pi_{t+12}^{w}]$	0.01	0.01***	0.01***	0.01**
	(0.01)	(0.00)	(0.00)	(0.01)
$\mathbb{E}_{it}^{\text{post}}[u_{t+12}]$	0.02*	0.03**	0.05***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
N	2,653	1,818	1,805	1,791
Sample	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes
F-stat for $\mathbb{E}_{it}^{\text{post}}[\pi_{t+12}$	8.26	8.29	12.74	8.94
F-stat for $\mathbb{E}_{it}^{\text{post}}[\pi_{t+12}^{w}]$	9.55	16.12	25.08	14.27
F-stat for $\mathbb{E}_{it}^{\text{post}}[u_{t+12}]$	19.73	13.05	17.13	14.19

Notes: This table presents the regression results for the effect of macroeconomic expectations on desired number of hours worked according to the following equation:

$$\begin{aligned}\mathbb{1}_{it}(\texttt{increase hours}) = &\beta_0 + \beta_1\mathbb{E}_{it}^{\text{post}}[\pi_{t+12}] + \beta_2\mathbb{E}_{it}^{\text{post}}[\pi_{t+12}^{w}] + \beta_3\mathbb{E}_{it}^{\text{post}}[u_{t+12}] + \theta\,\texttt{hours}_{it}\\ &+ \gamma_1\mathbb{E}_{it}^{\text{prior}}[\pi_{t+12}] + \gamma_2\mathbb{E}_{it}^{\text{prior}}[\pi_{t+12}^{w}] + \gamma_3\mathbb{E}_{it}^{\text{prior}}[u_{t+12}] + \mathbf{X}_{it}'\delta + \varepsilon_i\end{aligned}$$

We instrument the posterior expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, and the interaction of unemployment treatment dummies with prior expected unemployment rates. Highly numerate respondents in columns 3-4 are those who answered all the numerical competence check questions correctly. Heteroskedasticity-robust-standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for details about the treatment of outliers.

Overall, the results about offline labor supply preferences allow us to refine the puzzle arising from the analysis of the effect of wage growth expectations on the desired duration of employment on a specific MTurk project discussed in Section 3.4.1. These results suggest that workers do not decrease overall labor supply due to higher wage growth expectations, but instead reallocate it to other employers, either by working longer hours on day jobs or switching to other MTurk requesters who are likely offering higher-paid HITs.

3.5 Robustness to Alternative Specifications

This section discusses the robustness of the previously discussed results to alternative specifications. The main focus of the section is the analysis of the effect of information treatments through the framework of broad regime changes following Andrade et al. (2021) who provide evidence that what matters for households' decision-making is not the precise change in expectations but the broad regime changes. Additionally, we provide evidence about the robustness of the main results to alternative assumptions.

3.5.1 Information Treatment Effect on Broad Regime Changes in Expectations

Broad regime changes in expectations are indicator variables for the fact that respondents substantially switch their forecasts in response to information treatment (e.g., before treatment, respondents thought inflation would increase and after treatment, they thought it would decrease).

To evaluate the effect of information treatment on regime changes we estimate the following regression:[9]

$$\texttt{Regime Change}_i^Z = \beta_0 + \sum_{k\in\{\pi,\pi^w,u\}} \beta_{1,k}\texttt{treat}_i^k + \varepsilon_i, \;\; Z \in \{\pi, \pi^w, u\}, \tag{3.4}$$

where $\texttt{Regime Change}_i^Z$ denotes if a respondent i revises her *qualitative* assessment about variable Z *upwards*. For instance, if respondent i thinks that the overall price level will stay the same over the next 12 months, before the treatment, and changes this assessment so that she now thinks the overall price level will increase, after the treatment, then $\texttt{Regime Change}_i^\pi$ takes on the value of one. Similarly, if another respondent thinks that the overall price level will decrease over a year, before the treatment, but changes this assessment to "stay the same," or "increase," after the treatment, then $\texttt{Regime Change}_i^\pi$ is equal to one. It will take on the value of zero otherwise. We define $\texttt{Regime Change}_i^{\pi^w}$ similarly. Meanwhile, because unemployment rate expectations are elicited differently, we define $\texttt{Regime Change}_i^u$ as equal to one as long as respondents raise their unemployment expectations after the treatment and zero otherwise.

Table 3.8 shows the results. They paint the same picture as Table 3.3. First, columns 1-3 in table 3.8 show that when respondents receive either information about current CPI inflation rates or current hourly earnings inflation rates, they adjust their price inflation expectations upwards. Relative to those in the control group who received information about the air quality index in Seattle or COVID-19 vaccination rates, those in the CPI inflation treatment group are more likely to move to the higher CPI inflation rates regime by 3-6 percentage points. When respondents receive information about hourly earnings

[9] Appendix C.2.2 discusses results for an alternative specification that includes interactions of treatment dummies with prior expectations. They are qualitatively similar to the baseline results reported here.

inflation rates, they are more likely to change their price inflation expectation regimes upward by 4-7 percentage points.

Similarly, columns 4-6 of Table 3.8 show that respondents adjust their expected hourly earnings inflation rates upwards when they receive the relevant information. Relative to those in the control group, those in the CPI treatment group have a higher probability to move to a higher hourly earnings inflation regime by 2-4 percentage points. When they receive information about current hourly earnings inflation rates, they are more likely to move to a higher hourly earnings inflation regime by 12-16 percentage points relative to those in the control group.

Table 3.8: Information treatment effects on broad regime changes in forecast revisions

Dependent variable:	Price inflation ($Z = \pi$)			Wage inflation ($Z = \pi^w$)			Unemployment rate ($Z = u$)		
`Regime Change`$_i^Z$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
`treat_cpi`	0.06***	0.03*	0.04**	0.04**	0.02	0.02	0.08***	0.06***	0.04*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
`treat_wage`	0.07***	0.04***	0.05***	0.14***	0.12***	0.15***	0.01	-0.00	-0.05**
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
`treat_unemp`	0.01	0.02	0.04**	-0.03	-0.02	-0.00	-0.20***	-0.19***	-0.19***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	3903	3840	2810	3841	3766	2768	3694	3623	2637
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output from equation (3.4) for respondents in all control and treatment groups. The outcome variable is an indicator that respondents revised expectations of the variable in column header upward. For each outcome variable, the first column reports results without controls, the second column adds control variables, and the third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Lastly, columns 7-9 of Table 3.8 show that relative to those in the control group, those in the unemployment treatment group are *less* likely to move to higher unemployment rate regimes by 19-20 percentage points. In contrast, when they receive information about current CPI inflation rates, they tend to move to a higher unemployment rate regime. Compared to those in the control group, those in the CPI inflation treatment group are more likely to move to a higher unemployment rate regime by 4-8 percentage points. This is consistent with the result in Table 3.3 in Section 3.3 pointing to the stagflationary view of U.S. households.

3.5.2 Effect of Broad Regime Changes in Expectations on Labor Supply

Next, we discuss how broad regime changes affect online labor supply. We estimate regressions similar to equation (3.3) but now with dummy variables, $\texttt{Regime Change}^{Z}$ with $Z \in \{\pi, \pi^{w}, u\}$, denoting the broad regime changes before and after the information treatment, rather than the precise rate changes:

$$\Delta Y_{it}^{\texttt{post-prior}} = \beta_0 + \beta_1 \texttt{Regime Change}_i^{\pi} + \beta_2 \texttt{Regime Change}_i^{\pi^{w}} + \beta_3 \texttt{Regime Change}_i^{u} + \mathbf{X}_{it}'\delta + \varepsilon_i, \tag{3.5}$$

where $\Delta Y_{it} = \{\Delta \texttt{dur}_{it}^{\texttt{post}}, \Delta \overline{\texttt{rw}}_{it,t+\texttt{dur}_t}^{\texttt{post}}\}$ are changes in the desired duration of employment on our MTurk project (in month) and reservation wage per 10-minute monthly task. $\texttt{Regime Change}_i^{Z}$ is an indicator variable denoting if respondent i revises her *qualitative* assessment about a variable Z *upwards* defined in the same way as in Section 3.5.1. The first stage for this specification is summarized in Appendix C.2.2. The results reported in Table 3.9 are qualitatively similar to the baseline results in Section 3.4.

Table 3.9: Effects of regime changes in expectations on MTurk labor supply

	Desired Duration (in months)				Reservation Wages			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\texttt{Regime}^{\pi}$	-0.03	0.03	-0.00	0.10	-0.13**	-0.02	-0.04	-0.08
	(0.10)	(0.11)	(0.12)	(0.13)	(0.07)	(0.07)	(0.09)	(0.13)
$\texttt{Regime}^{\pi^{w}}$	-0.17	-0.25**	-0.22**	-0.24**	0.17***	0.20***	0.12**	0.19***
	(0.10)	(0.10)	(0.10)	(0.10)	(0.06)	(0.06)	(0.05)	(0.06)
$\texttt{Regime}^{u}$	0.47***	0.46***	0.41***	0.35***	-0.14**	-0.03	-0.04	0.04
	(0.10)	(0.10)	(0.10)	(0.10)	(0.05)	(0.06)	(0.05)	(0.06)
N	3,127	3,088	2,203	2,158	3,051	3,005	2,118	2,080
Sample	All	All	Numerate	Numerate	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-stat for $\texttt{Regime}^{\pi}$	29.73	26.94	16.81	15.34	26.70	22.47	13.71	10.43
F-stat for $\texttt{Regime}^{\pi^{w}}$	36.40	34.60	24.69	27.38	32.33	34.31	21.38	22.49
F-stat for $\texttt{Regime}^{u}$	41.77	30.43	40.57	30.67	36.08	24.92	37.25	26.94

Notes: This table presents the regression output to estimate the effects of broad regime changes in expectations on MTurk labor supply for equation (3.5). We instrument the regime changes in expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, and the interaction of unemployment treatment dummies with prior expected unemployment rates. Highly numerate respondents are those who answered all the numerical competence check questions correctly. Heteroskedasticity-robust-standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for details about the treatment of outliers.

Columns 1-4 of Table 3.9 report the results for the desired duration of employment. As respondents increase their expected unemployment rates, they increase their desired duration of employment. Broad changes in the price inflation regime do not affect the desired duration of employment, but the broad changes in hourly earnings inflation do. As respondents move from no change in hourly earnings to an increase in hourly earnings or from a decrease in hourly earnings to no change in hourly earnings and/or increase in hourly earnings, they decrease their desired duration of employment on our MTurk project. However, they are likely to switch to other employers rather than decrease the overall labor supply (see Section 3.4.3).

Columns 5-8 in Table 3.9 report the results for MTurk reservation wages. They are qualitatively similar to those in Table 3.5, but some coefficients are not statistically significant given that there is less variation in endogenous variables. As respondents revise their broad regime about hourly earnings inflation expectation upwards, they increase their reservation wages. The upward revision of price inflation expectations, however, is associated with the decrease in reservation wages, but the coefficients are not statistically significant. Similarly, the upward forecast revisions of unemployment rates are associated with lower reservation wages in most specifications but the results are not statistically significant.

3.5.3 Additional Robustness Checks

This section provides evidence about the robustness of the results to adjustment of p-values for multiple hypothesis testing, alternative instruments sets, and dependent variable.

Adjustment of $p-$values for Multiple Hypothesis Testing. To address the concern that having three endogenous variables in our preferred specification biases standard errors and, thus, invalidates hypothesis testing, in this section we discuss the results with adjusted p-values. When estimating equation (3.3), we are interested in six parameter values. The regression coefficients on the forecast revisions in price and wage inflation rates, and unemployment rates with two dependent variables: the desired duration of employment and the reservation wages. To minimize the likelihood of false rejections with multiple hypothesis testing, we use Westfall-Young stepdown adjusted p-values using `wyoung` command in STATA. This procedure controls the familywise error rate (FWER) and allows for dependence amongst p-values. The results with adjusted p-values are reported in Appendix C.4.1. They are similar to the main results about the effect of expectations on labor supply both in terms of continuous expectations revisions and discrete regime changes.

Alternative Instruments. The main specifications considered in Sections 3.4 and 3.5.2 instruments three endogenous expectations variables with a set of information treatment dummies, the interaction of prior price inflation expectations with the CPI treatment dummy and hourly earnings treatment dummies, the interaction of prior hourly earnings inflation expectations with the CPI treatment dummy and hourly earnings treatment dummy, and the interaction of prior unemployment expectations with unemployment treatment. Alternatively, we could include additional interaction terms with the unemployment treatment dummy as well as prior and posterior unemployment expectations. The results reported in Appendix C.4.2 for the alternative set of instruments are similar to the baseline IV results. However, a baseline specification is preferred because it produces a stronger first stage by excluding weaker instruments.

Alternative Dependent Variable. The results discussed so far focus on the relationship between the *changes* in labor supply on revisions expectations. Alternatively, we could modify the regression specification to have the *level* of labor supply as an outcome variable, i.e., the post-treatment desired number of months worked and post-treatment reservation wage, and introduce the pre-treatment version of the outcome as a control. The results in Appendix C.4.3 show that such a specification change does not affect results much relative to the baseline. The main differences are that desired duration of employment does not significantly respond to expected unemployment rates and wage growth rates as before. However, conclusions about the positive effect of wage growth and the negative effect of price inflation expectations on reservation wages are still valid.

3.6 Discussion and Conclusions

We study how changes in macroeconomic expectations affect labor supply preferences by conducting an experiment in an online labor market. To this end, we generate exogenous variation in subjective expectations about price inflation, wage inflation, and unemployment rates by randomizing information treatments. We then use the resulting exogenous variation in expectations to study how it affects MTurk workers' reservation wages and the desired employment duration. Our results provide the first direct causal evidence about the effect of inflation expectations on labor supply and suggest that the risks of wage-price spirals are limited in the current high inflation setting.

First, we show that respondents significantly revise their macroeconomic expectations when provided with relevant information. Importantly, in response to a signal about one variable (e.g., unemployment rate) respondents revise multiple expectations *jointly*. When workers revise unemployment or wage growth expectations, they tend to revise price inflation expectations in the same direction. While inflation expectations are the most responsive to signals about other variables, unemployment expectations are mostly responsive to signals about unemployment. Wage inflation expectations tend to comove with price inflation expectations.

Next, exploiting the resulting variation in macroeconomic expectations, we document several results about the effect of expectations on labor supply. First, we find that respondents decrease their desired duration of employment on our MTurk project in response to higher wage inflation expectations. Second, we document that desired duration of employment on our project increases in unemployment expectations. It suggests that when workers become more pessimistic about the aggregate labor market situation, they are more willing to secure long-term employment with a specific employer. Third, we find that desired duration of employment does not significantly respond to changes in price inflation expectations. Fourth, we document that higher wage inflation expectations increase reservation wages. Fifth, higher price inflation expectations appear to *decrease* reservation wages whereas higher unemployment expectations do not significantly affect reservation wages.

The fact that higher wage inflation rates decrease the number of periods workers would like to commit to work with us does not necessarily extrapolate to other online employers (MTurk requesters) or offline employment. We cannot rule out the possibility that rather than commit to working with one employer at pre-determined pay, with higher expected wage inflation rates, workers might want to switch to other employers looking for higher-paid HITs, without changing the total hours worked. Based on supplementary evidence about offline labor market preferences, we interpret this result as evidence of labor supply reallocation across employers rather than a decrease in total labor supply.

The result that wage and price inflation expectations affect reservation wages in opposite directions has important implications for understanding how households interpret inflation. This interpretation matters for the likelihood of wage-price spirals. The fact that reservation wages are increasing in wage growth expectations is not surprising. However, the fact that workers are willing to accept work at lower pay due to an increase in inflation expectations, rather than demanding additional compensation to restore the purchasing power of their income, is surprising. This result implies that the response of labor supply to inflation mitigates the threat of wage-price spirals. From the perspective of a search-theoretic model (e.g., Rogerson et al., 2005), the observed response to inflation expectations shock is consistent with households interpreting an increase in price inflation expectations as a signal about the deterioration of outside options, which induces them to reduce reservation wages and duration for job search/unemployment. The response to an increase in wage inflation expectations is similar to the reaction to an increase in outside options.

There is additional evidence that points to the fact that households interpret an inflation increase as a cautionary sign. When analyzing revision of expectations in response to a randomized information provision, we find that households associate higher inflation rates with higher unemployment rates. Respondents tend to increase their expected unemployment rates when provided with the current inflation rates. This is consistent with the evidence in the literature that U.S. households tend to exhibit the stagflationary view (see Kamdar, 2018; Binder, 2020). This result suggests that the first chain of wage-price spirals could be partially muted with higher expected unemployment rates.

Our results are based on the experiments conducted in an online labor market, Amazon MTurk, which has distinctive features compared to offline labor markets. Online labor markets, in particular, feature much greater flexibility. It is much easier for workers to adjust their labor supply in online labor markets than in offline labor markets. Because MTurk workers are much more flexible, they represent those who are on the margin of adjustment and about whom policymakers care the most. Moreover, because in the follow-up surveys, we offer workers employment on the terms provided by them, we were able to capture the "actual" labor supply preferences as opposed to hypothetical preferences based on hypothetical questions only. At the same time, however, because of the distinctive features of online labor markets, offline labor supply responses could be different from our results to some extent. Due to the inflexibility, we might not be able to observe responses to the same degree. Because most workers use offline labor markets as their primary income source, their labor supply responses could be much larger. How much offline responses are different from online responses is left for our future work.

Bibliography

Acemoglu, D., Hassan, T. A., and Robinson, J. A. (2011). Social structure and development: A legacy of the Holocaust in Russia. *The Quarterly Journal of Economics*, 126(2):895–946.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Adler, M. A. (2002). German unification as a turning point in east german women's life course: Biographical changes in work and family roles. *Sex Roles*, 47:83–98.

Aguiar, M. and Hurst, E. (2016). The macroeconomics of time allocation. In *Handbook of Macroeconomics*, volume 2, pages 203–253. Elsevier.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Albanesi, S. and Olivetti, C. (2009). Home production, market production and the gender wage gap: Incentives and expectations. *Review of Economic Dynamics*, 12(1):80–107.

Alesina, A. and Fuchs-Schündeln, N. (2007). Goodbye Lenin (or not?): The effect of communism on people's preferences. *American Economic Review*, 97(4):1507–1528.

Algan, Y. and Cahuc, P. (2014). Trust, growth, and well-being: New evidence and policy implications. In *Handbook of Economic Growth*, volume 2, pages 49–120. Elsevier.

Allen, R. C. (2021). *Farm to factory*. Princeton University Press.

Almqvist, A.-L. and Duvander, A.-Z. (2014). Changes in gender equality? Swedish fathers' parental leave, division of childcare and housework. *Journal of Family Studies*, 20(1):19–27.

Andrade, P., Gautier, E., and Mengus, E. (2021). What matters in households' inflation expectations? Banque de France Working Paper.

Angrist, J. D. and Frandsen, B. (2022). Machine labor. *Journal of Labor Economics*, 40(S1):S97–S140.

Applebaum, A. (2017). *Red Famine: Stalin's War on Ukraine*. Doubleday.

Armantier, O., Nelson, S., Topa, G., Van der Klaauw, W., and Zafar, B. (2016). The price is right: Updating inflation expectations in a randomized price information experiment. *The Review of Economics and Statistics*, 98(3):503–523.

Armona, L., Fuster, A., and Zafar, B. (2019). Home price expectations and behaviour: Evidence from a randomized information experiment. *The Review of Economic Studies*, 86(4):1371–1410.

Asatkina, O. (1935). *Narodne hospodarstvo USRR: (stat.dovid.) / Upr. nar.hosp. oblik USRR [National economy of the Ukrainian SSR: (statistical report): Governance of national accounting of USSR]*. Kyiv: National economy and accounting.

Baranskaya, N. and Lehrman, E. (1974). A week like any other week. *The Massachusetts Review*, 15(4):657–703.

Bauernschuster, S., Falck, O., Gold, R., and Heblich, S. (2012). The shadows of the socialist past: Lack of self-reliance hinders entrepreneurship. *European Journal of Political Economy*, 28(4):485–497.

Baxter, J. and Tai, T.-o. (2016). Inequalities in unpaid work: A cross-national comparison. *Handbook on Well-being of Working Women*, pages 653–671.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal*, 75(299):493–517.

Becker, G. S. (1991). *A treatise on the family: Enlarged edition*. Harvard University Press.

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014a). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014b). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.

Belot, M., Kircher, P., and Muller, P. (2022). How wage announcements affect job search—a field experiment. *American Economic Journal: Macroeconomics*, 14(4):1–67.

Bertrand, M. (2020). Gender in the twenty-first century. *AEA Papers and Proceedings*, 110:1–24.

Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130(2):571–614.

Besley, D., Kuh, E., and Welsch, R. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley and Sons.

Bianchi, S. M., Milkie, M. A., Sayer, L. C., and Robinson, J. P. (2000). Is anyone doing the housework? Trends in the gender division of household labor. *Social Forces*, 79(1):191–228.

Binder, C. (2020). Coronavirus fears and macroeconomic expectations. *The Review of Economics and Statistics*, 102(4):721–730.

Binder, C. and Kim, G. (2020). Learning-through-survey in inflation expectations. *American Economic Journal: Macroeconomics*.

Bittman, M., England, P., Sayer, L., Folbre, N., and Matheson, G. (2003). When does gender trump money? Bargaining and time in household work. *American Journal of Sociology*, 109(1):186–214.

Blau, F. D., Kahn, L. M., Comey, M., Eng, A., Meyerhofer, P., and Willén, A. (2020). Culture and gender allocation of tasks: Source country characteristics and the division of non-market work among US immigrants. *Review of Economics of the Household*, 18:907–958.

Bluhm, R. and McCord, G. C. (2022). What can we learn from nighttime lights for small geographies? Measurement errors and heterogeneous elasticities. *Remote Sensing*, 14(5):1190.

Boissay, F., De Fiore, F., Igan, D., Tejada, A. P., Rees, D., et al. (2022). Are major advanced economies on the verge of a wage-price spiral? BIS Bulletin No. 53, Bank for International Settlements.

Bottan, N. L. and Perez-Truglia, R. (2020). Betting on the house: Subjective expectations and market choices. Working Paper 27412, National Bureau of Economic Research.

Brainerd, E. (2000). Women in transition: Changes in gender wage differentials in Eastern Europe and the former Soviet Union. *ILR Review*, 54(1):138–162.

Brines, J. (1994). Economic dependency, gender, and the division of labor at home. *American Journal of Sociology*, 100(3):652–688.

Bystydzienski, J. M. (1989). Women and socialism: A comparative study of women in Poland and the USSR. *Signs: Journal of Women in Culture and Society*, 14(3):668–684.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.

Campa, P. and Serafinelli, M. (2019). Politico-economic regimes and attitudes: Female workers under state socialism. *The Review of Economics and Statistics*, 101(2):233–248.

Cavallo, A., Cruces, G., and Perez-Truglia, R. (2017). Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics*, 9(3):1–35.

Chatterjee, C. (2001). Soviet heroines and the language of modernity, 1930–39. In *Women in the Stalin Era*, pages 49–68. Springer.

Chen, X. and Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21):8589–8594.

Chen, Y. and Yang, D. (2019). Historical traumas and the roots of political distrust: Political inference from the Great Chinese Famine. Working paper.

Cheremukhin, A., Golosov, M., Guriev, S., and Tsyvinski, A. (2017). The industrialization and economic development of Russia through the lens of a neoclassical growth model. *The Review of Economic Studies*, 84(2):613–649.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., and Van Rooij, M. (2023). How does consumption respond to news about inflation? Field evidence from a randomized control trial. *American Economic Journal: Macroeconomics*, forthcoming.

Coibion, O., Gorodnichenko, Y., Knotek, E. S., and Schoenle, R. (2022a). Average inflation targeting and household expectations. *Journal of Political Economy–Macroeconomics*, forthcoming.

Coibion, O., Gorodnichenko, Y., Kumar, S., and Ryngaert, J. (2021). Do you know that I know that you Know…? Higher-order beliefs in survey data. *The Quarterly Journal of Economics*, 136(3):1387–1446.

Coibion, O., Gorodnichenko, Y., and Weber, M. (2022b). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6):1537–1584.

Colella, F., Lalive, R., Sakalli, S. O., and Thoenig, M. (2019). Inference with arbitrary clustering. Discussion paper No. 12584, IZA.

Coltrane, S. (2000). Research on household labor: Modeling and measuring the social embeddedness of routine family work. *Journal of Marriage and Family*, 62(4):1208–1233.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Coverman, S. (1985). Explaining husbands'participation in domestic labor. *Sociological Quarterly*, 26(1):81–97.

Davies, R. and Wheatcroft, S. (2016). *The years of hunger: Soviet agriculture, 1931–1933*. Springer.

Davies, R. W. et al. (1998). *Crisis and progress in the Soviet economy, 1931-1933*. Macmillan Press.

Davis, S. N. and Greenstein, T. N. (2009). Gender ideology: Components, predictors, and consequences. *Annual Review of Sociology*, 35:87–105.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica*, 78(6):1863–1903.

Dell, M., Jones, B. F., and Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, 4(3):66–95.

Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature*, 52(3):740–98.

Doepke, M. and Zilibotti, F. (2014). Culture, entrepreneurship, and growth. In *Handbook of economic growth*, volume 2, pages 1–48. Elsevier.

Doll, C. N., Muller, J.-P., and Elvidge, C. D. (2000). Night-time imagery as a tool for global mapping of socioeconomic parameters and greenhouse gas emissions. *Ambio*, pages 157–162.

Dvorkin, M. A. and Shell, H. (2015). A cross-country comparison of labor force participation. *Economic Synopses*, (17).

EBRD (1995). Investment and enterprise development. Transition report, European Bank for Reconstruction and Development.

Einhorn, B. (1993). *Cinderella goes to market*. Verso London.

Esping-Andersen, G. (1989). The three political economies of the welfare state. *Canadian Review of Sociology*, 26(1):10–36.

Fritsch, M., Bublitz, E., Sorgner, A., and Wyrwich, M. (2014). How much of a socialist legacy? The re-emergence of entrepreneurship in the East German transformation to a market economy. *Small Business Economics*, 43(2):427–446.

Funk, N. and Mueller, M. (1993). *Gender politics and post-communism: Reflections from Eastern Europe and the former Soviet Union*. Routledge.

Fuwa, M. (2004). Macro-level gender inequality and the division of household labor in 22 countries. *American Sociological Review*, 69(6):751–767.

Galenson, A. et al. (2004). *Economies in Transition: An OED Evaluation of World Bank Assistance*. World Bank Publications.

Galligan, Y., Clavero, S., and Calloni, M. (2008). *Gender politics and democracy in post-socialist Europe*. Barbara Budrich.

Geist, C. (2009). One Germany, two worlds of housework? Examining employed single and partnered women in the decade after unification. *Journal of Comparative Family Studies*, 40(3):415–437.

Ghodsee, K. (2018). *Why women have better sex under socialism: And other arguments for economic independence*. Random House.

Ghodsee, K. and Mead, J. (2018). What has socialism ever done for women? *Catalyst*, 2(2):101–133.

Ghosh, T., L Powell, R., D Elvidge, C., E Baugh, K., C Sutton, P., and Anderson, S. (2010). Shedding light on the global distribution of economic activity. *The Open Geography Journal*, 3(1).

Gilchrist, D. S. and Sands, E. G. (2016). Something to talk about: Social spillovers in movie consumption. *Journal of Political Economy*, 124(5):1339–1382.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4):1091–1119.

Goldman, W. (2001). *Babas* at the bench: Gender conflict in Soviet industry in the 1930s. In *Women in the Stalin Era*, pages 69–88. Springer.

Gorodnichenko, Y. and Roland, G. (2011). Which dimensions of culture matter for long-run growth? *American Economic Review*, 101(3):492–98.

Gregory, P. R. (2014). *Before command: an economic history of Russia from emancipation to the first five-year*, volume 216. Princeton University Press.

Grosfeld, I., Rodnyansky, A., and Zhuravskaya, E. (2013). Persistent antimarket culture: A legacy of the Pale of Settlement after the Holocaust. *American Economic Journal: Economic Policy*, 5(3):189–226.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, 61(1):3–40.

Hajdini, I., Knotek, E. S., Leer, J., Pedemonte, M., Rich, R. W., and Schoenle, R. (2022a). Indirect consumer inflation expectations: Theory and evidence. *Economic Commentary*, 2022-03.

Hajdini, I., Knotek, E. S., Leer, J., Pedemonte, M., Rich, R. W., and Schoenle, R. (2022b). Low passthrough from inflation expectations to income growth expectations: Why people dislike inflation. Working paper.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Holodomor Museum (2021). Worldwide recognition of the Holodomor as genocide. https://holodomormuseum.org.ua/en/recognition-of-holodomor-as-genocide-in-the-world. [Accessed August 21, 2022].

Hrynevych, L. (2013). *Holod 1928-1929 rr. u Radianskiy Ukraiini [The famine of 1928-1929 in Soviet Ukraine]*. Kyiv: Institute of History of Ukraine, The National Academy of Science of Ukraine.

Hunter, H. (1973). The overambitious first Soviet five-year plan. *Slavic Review*, 32(2):237–257.

Ishizawa, O. A., Miranda, J. J., and Strobl, E. (2019). The impact of Hurricane strikes on short-term local economic activity: Evidence from nightlight images in the Dominican Republic. *International Journal of Disaster Risk Science*, 10(3):362–370.

Jann, B. et al. (2008). A Stata implementation of the Blinder-Oaxaca decomposition. *Stata Journal*, 8(4):453–479.

Kamdar, R. (2018). The inattentive consumer: Sentiment and expectations. Working paper.

Kan, M. Y., Sullivan, O., and Gershuny, J. (2011). Gender convergence in domestic work: Discerning the effects of interactional and institutional barriers from large-scale data. *Sociology*, 45(2):234–251.

Kandil, M. M. E. (2003). The wage-price spiral: Industrial country evidence and implications. IMF Working Paper.

Kessler, G. (2001). The passport system and state control over population flows in the Soviet Union, 1932-1940. *Cahiers du monde russe. Russie-Empire russe-Union soviétique et États indépendants*, 42(42/2-4):477–504.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126.

Kranz, S. (2005). Women's role in the German Democratic Republic and the state's policy toward women. *Journal of International Women's Studies*, 7(1):69–83.

LaFont, S. (2001). One step forward, two steps back: Women in the post-communist states. *Communist and Post-Communist Studies*, 34(2):203–220.

Lemkin, R. (1953). Soviet genocide in the Ukraine. In *Holodomor: Reflections on the Great Famine of 1932–1933 in Soviet Ukraine (2008), edited by Luciuk, L.Y. and Grekul, L.*, chapter 10, pages 266–290. The Kashtan Press.

Lenin, V. I. (1919). Soviet power and the status of women. *First published: Pravda No. 249, November 6, 1919*.

Lewin, M. (1966). Who was the Soviet kulak? *Europe-Asia Studies*, 18(2):189–212.

Lind, J. T. (2019). Spurious weather effects. *Journal of Regional Science*, 59(2):322–354.

Lippmann, Q., Georgieff, A., and Senik, C. (2020). Undoing gender with institutions: Lessons from the German division and reunification. *The Economic Journal*, 130(629):1445–1470.

Lupu, N. and Peisakhin, L. (2017). The legacy of political violence across generations. *American Journal of Political Science*, 61(4):836–851.

Markevich, A., Naumenko, N., and Qian, N. (2021). The political-economic causes of the Soviet Great Famine, 1932–33. Working Paper 29089, National Bureau of Economic Research.

Matsuura, K. and Willmott, C. (2014). Terrestrial air temperature and precipitation: 1900–2014 gridded monthly time series (version 4.01). *National Oceanic and Atmospheric Administration, https://www. esrl. noaa. gov/psd/data/gridded/data. UDel_ AirT_Precip. html*.

Maynes, C. (2021). Russia lifts Soviet-era rules on what jobs women can do. In *NPR Morning Edition*, Podcast audio.

Mellander, C., Lobo, J., Stolarick, K., and Matheson, Z. (2015). Night-time light data: A good proxy measure for economic activity? *PloS One*, 10(10):e0139779.

Meng, X., Qian, N., and Yared, P. (2015). The institutional causes of China's Great Famine, 1959–1961. *The Review of Economic Studies*, 82(4):1568–1611.

Mikucka, M. (2009). Division of household labor between spouses: How do Central and Eastern Europe differ from the West? *International Journal of Sociology*, 39(1):76–94.

Mishchenko, T. (2011). Problemma socialnogo vremeni v sisteme tsennostnych orientatsiy sovetskikh zhenshchin v 1960-80h godah: Gosudarstvennaya politika i individual'nyi vybor [The problem of social time in the value orientations of Soviet women in the 1960-1980s: Public policy and individual choice]. *Society. Communication. Education*, (3(131)):160–164.

Motyl, A. J. (2010). Deleting the Holodomor: Ukraine unmakes itself. *World Affairs*, 173(3):25–33.

Mykhaylychenko, G. and Shatalina, Y. (1992). *Collectivization and famine in Ukraine: 1929-1933: Collection of materials and documents [Koleklyvizatsia i holod na Ukraini: 1929-1933: Zbirnyk materialiv i dokumentiv]*. Kyiv: Naukova Dumka.

Naumenko, N. (2019). Economic consequences of the 1933 Soviet famine. Working paper, George Mason University.

Naumenko, N. (2021). The political economy of famine: The Ukrainian famine of 1933. *The Journal of Economic History*, 81(1):156–197.

Papakin, H. (2010). The list of villages/village councils and collective farms of the Ukrainian SSR blacklisted in 1932-1933. Institute of History of Ukraine, The National Academy of Science of Ukraine.

Papakin, H. (2013). *"Chorna doshka": Antyselianski represii (1932-1933) ["Black Board": Antipeasant Repressions (1932-1933)]*. Kyiv: Institute of History of Ukraine, The National Academy of Science of Ukraine.

Prakash, A., Shukla, A. K., Bhowmick, C., and Beyer, R. C. M. (2019). Night-time luminosity: Does it brighten understanding of economic activity in India. *Reserve Bank of India Occasional Papers*, 40(1):1–24.

Rogerson, R., Shimer, R., and Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4):959–988.

Roland, G. (2002). The political economy of transition. *Journal of Economic Perspectives*, 16(1):29–50.

Rozenas, A., Schutte, S., and Zhukov, Y. (2017). The political legacy of violence: The long-term impact of Stalin's repression in Ukraine. *The Journal of Politics*, 79(4):1147–1161.

Rozenas, A. and Zhukov, Y. (2019). Mass repression and political loyalty: Evidence from Stalin's' terror by hunger. *American Political Science Review*.

Rozovyk, O. (2020). Resettlement processes in the Ukrainian SSR during the Holodomor (1932–34). *Communist and Post-Communist Studies*, 53(1):80–99.

Rudnytskyi, O., Levchuk, N., Wolowyna, O., Shevchuk, P., and Kovbasiuk, A. (2015). Demography of a man-made human catastrophe: The case of massive famine in Ukraine 1932–1933. *Canadian Studies in Population*, 42(1-2):53–80.

Runfola, D., Anderson, A., Baier, H., Crittenden, M., Dowker, E., Fuhrig, S., Goodman, S., Grimsley, G., Layko, R., Melville, G., et al. (2020). geoBoundaries: A global database of political administrative boundaries. *PLoS One*, 15(4):e0231866.

Saxonberg, S. (2014). *Gendering family policies in post-communist Europe: A historical-institutional analysis*. Springer.

Scholz, E., Jutz, R., Edlund, J., Öun, I., and Braun, M. (2014). ISSP 2012 Family and Changing Gender Roles IV: Questionnaire Development.

Shandra, A. (2018). See which countries recognize Ukraine's Holodomor famine as genocide on an interactive map. Euromaidan Press, https://euromaidanpress.com/2018/11/24/see-which-countries-recognize-ukraines-holodomor-famine-as-genocide-on-an-interactive-map. [Accessed August 21, 2022].

Siminski, P. and Yetsenga, R. (2022). Specialization, comparative advantage, and the sexual division of labor. *Journal of Labor Economics*, 40(4):851–887.

Snyder, T. (2010). *Bloodlands: Europe between Hitler and Stalin*. Basic Books (AZ).

Sonin, K. (2013). The end of economic transition. *Economics of Transition*, 21(1):1–10.

Stasiuk, O., Lukianets, K., Starovoit, S., and Baikenich, D. (2021). *"Chorni doshky" Ukraiiny. Chernihivs'ka oblast / Zbirnyk dokumentiv i materialiv. ["Black boards" of Ukraine. Chernihiv oblast / Documents and materials]*. Kyiv: Marko Melnyk publishing.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4):518–529.

Treas, J. and Tai, T. (2016). Gender inequality in housework across 20 European nations: Lessons from gender stratification theories. *Sex Roles*, 74:495–511.

Ukhova, D. (2020). Gender division of domestic labor in post-socialist Europe (1994–2012): Test of class gradients hypothesis. *Social Inclusion*, 8(4):23–34.

Viola, L. (2000). The peasants' kulak: Social identities and moral economy in the Soviet countryside in the 1920s. *Canadian Slavonic Papers*, 42(4):431–460.

Walden, J. and Zhukov, Y. M. (2020). Historical legacies of political violence. In *Oxford Research Encyclopedia of Politics*.

Wolowyna, O., Plokhy, S., Levchuk, N., Rudnytskyi, O., Shevchuk, P., and Kovbasiuk, A. (2016). Regional variations of 1932–34 famine losses in Ukraine. *Canadian Studies in Population*, 43(3-4):175–202.

World Bank (2002). Ukraine country brief.

Yakubova, L. (2011). *Povsiakdenne zhyttia etnichnyh menshyn Radianskoi Ukrainy u mizhvoenny dobu [Day-to-day life of ethnic minorities in Soviet Ukraine during the interwar period]*. Institute of History of Ukraine.

Appendix A

Appendix for Chapter 1

A.1 Additional Figures and Tables

This Appendix section provides additional figures and tables relevant for understanding historical context, and the effect of blacklisting.

Historical context

- Figure A.1 shows that acceleration of collectivization in early 1930s was accompanied by a rapid drop of livestock in Ukraine.
- Figure A.2 shows spatial variation in collectivization rate in Ukraine in the end of 1932. Districts in the main grain producing areas in the south have the highest collectivization rates as prioritized by the authorities.
- Figure A.3 shows the grain harvest in Ukraine in 1923-1933 with breakdown for rye and wheat. The harvest declined between 1930-1932 but recovered in 1933.
- Figure A.4 shows district-level deviation from grain procurement plain in 1930-1933. The number of districts that met or exceeded the plan declined over time.
- Figure A.5 shows the distribution of spatial distribution of rural famine losses on the district level.

Effect of blacklisting

- Figure A.6 is the schematic map for channels through which local weather shocks affect long-term outcomes to complement identification strategy discussed in Section 1.4.2.
- Table A.1 provides first stage results for alternative sets of instruments. A preferred set of instruments is in column 4.

- Table A.2 compares the characteristics of districts with and without blacklisted communities
- Table A.3 reports the coefficients on control variables omitted in Table 1.4.
- Tables A.4-A.5 report the effect of blacklisting on population size and ethnic composition to supplement the results in Section 1.7.3.

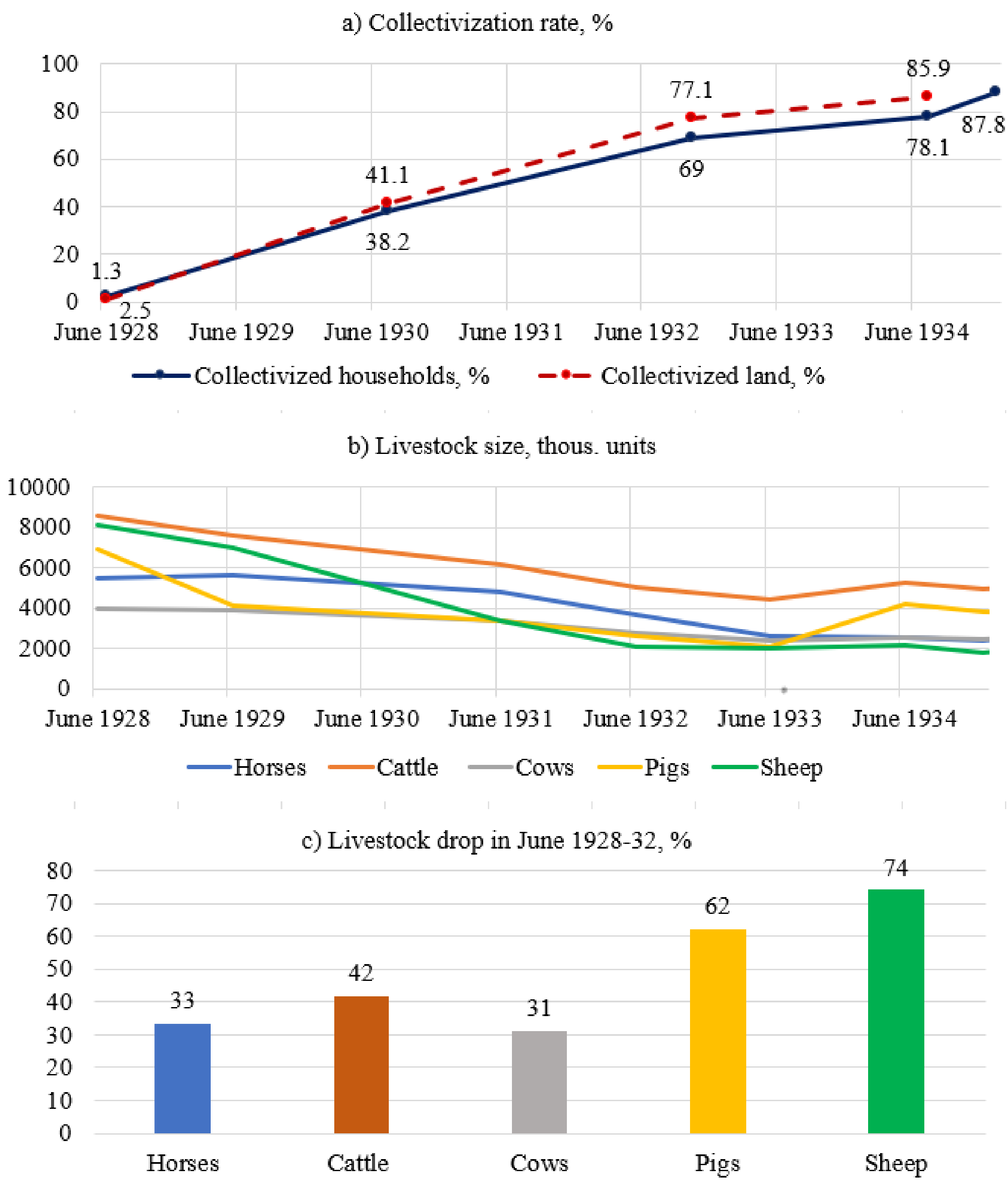


Figure A.1: Collectivization pace and livestock drop

Source: Constructed by author based on data from (Asatkina, 1935)

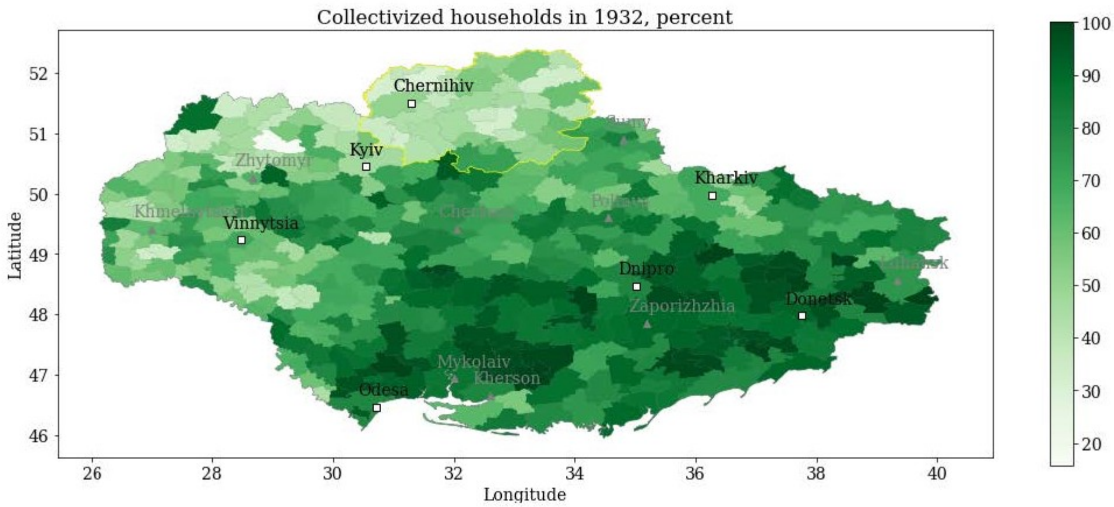


Figure A.2: Share of collectivized households in the end of 1932

Source: Based on HURI Famine Web Map data.

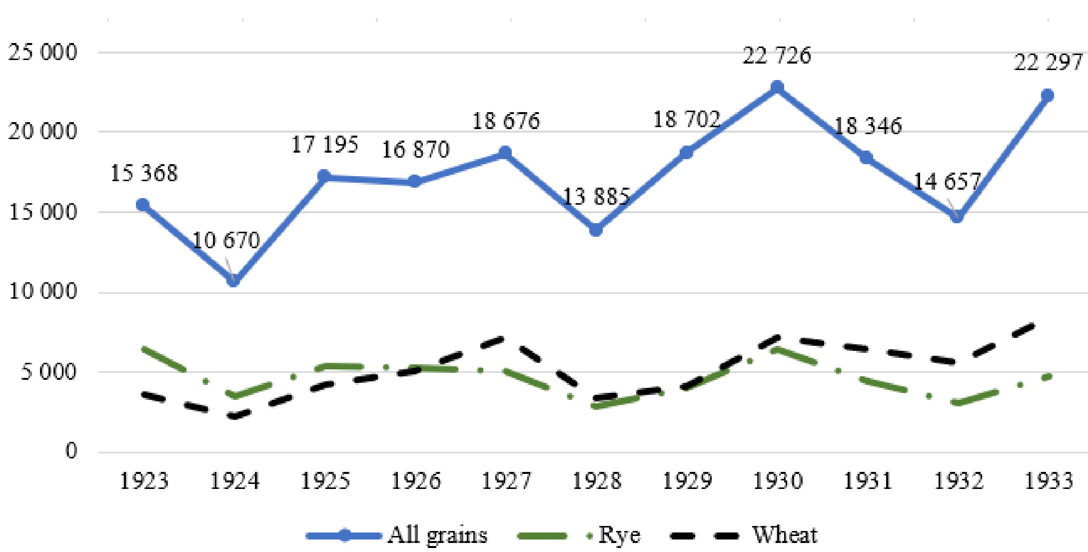


Figure A.3: Grain harvest, thous. tons

Source: Constructed by author based on data from Asatkina (1935). 1928-29 and 1932-33 years of famine. The numbers should be interpreted with caution due to difficulties of harvest measurement and possibility of misreporting.

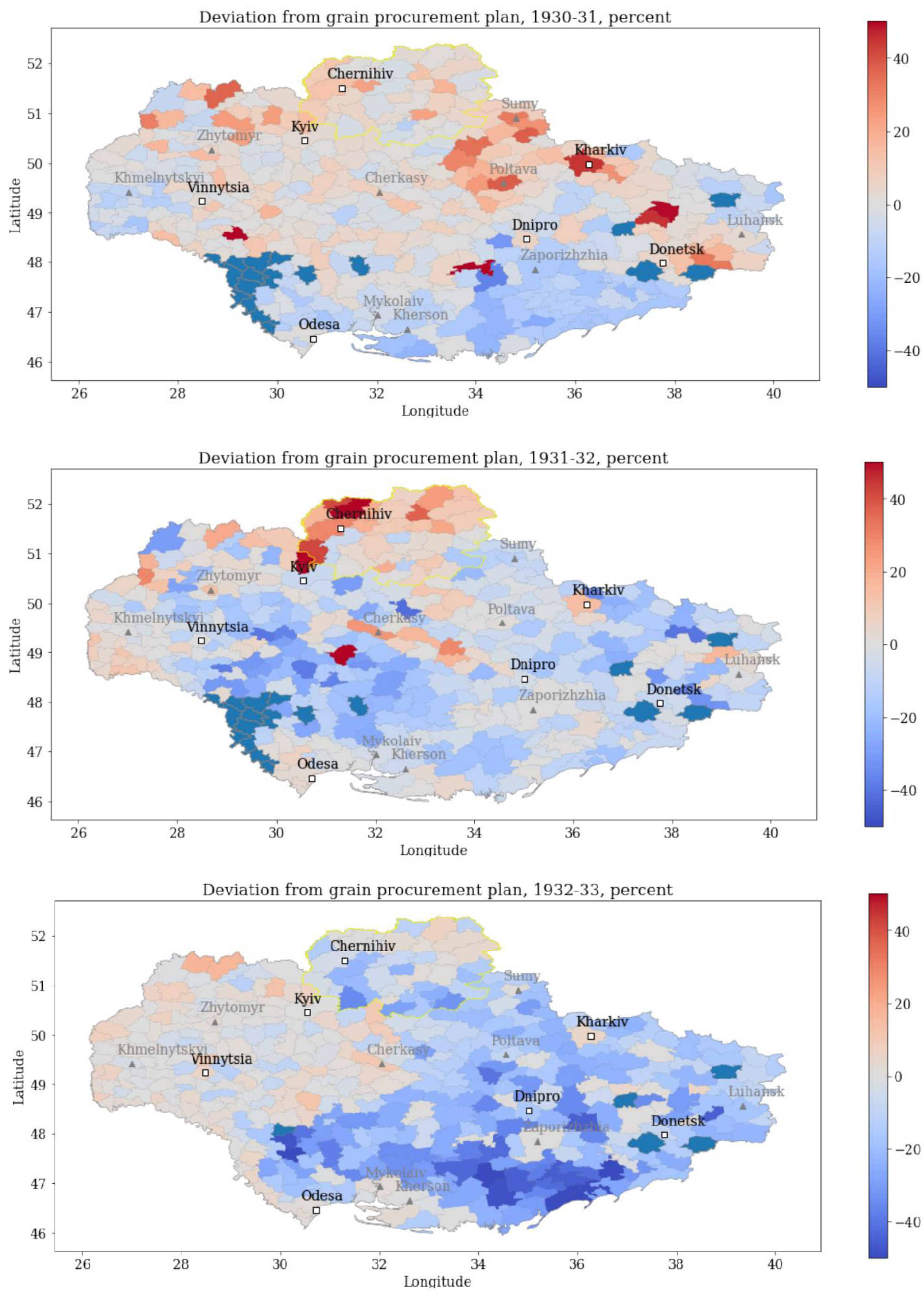


Figure A.4: Deviation from procurement plan in 1930-1932

Source: Based on Famine Web Map data. Positive values indicate exceeding the plan.

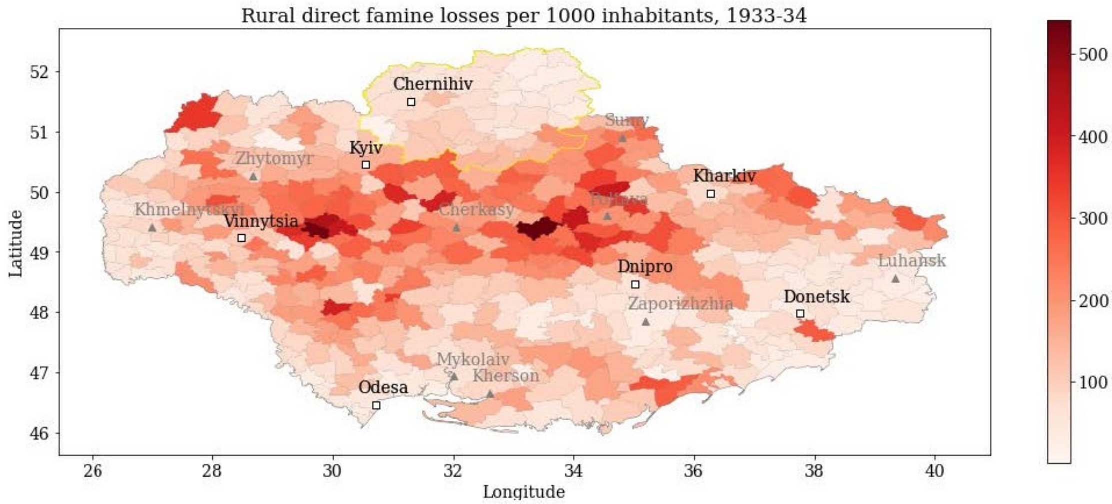


Figure A.5: Rural famine losses per 1,000 inhabitants, 1933-34

Source: Based on HURI Famine Web Map data.

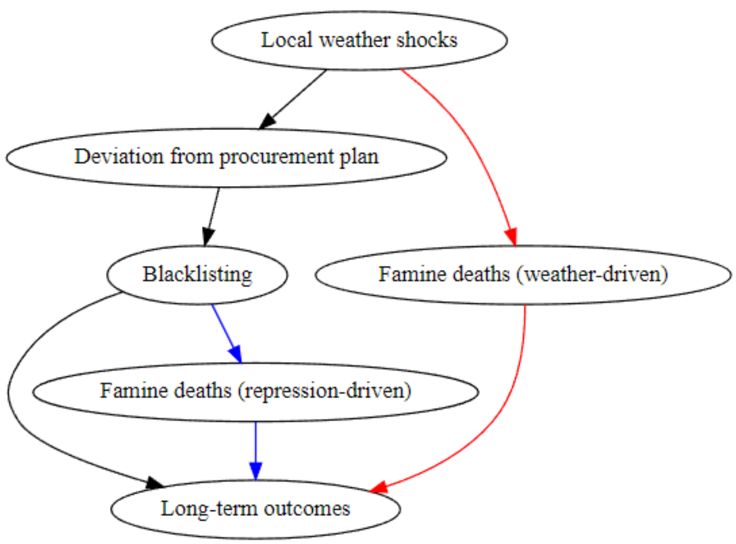


Figure A.6: Schematic map for the long-term effect of blacklisting

Notes: Black lines illustrate the direct channel for the effect of blacklisting on long-term outcomes and blue lines – indirect channel, intermediated by the famine mortality inflicted by blacklisting. Red lines denote an alternative channel of how weather shocks may affect long-term outcomes (through famine mortality unrelated to blacklisting). While this alternative channel is valid, it is arguably small due to evidence that weather was a less important determinant of the famine than the collectivist policies (Naumenko, 2021). Therefore, omitting this effect should not bias the results substantially.

Table A.1: First stage for various sets of instruments

	(1) DevPlan_pl	(2) DevPlan_ad,cv	(3) DevPlan_pl	(4) DevPlan_ad	(5) DevPlan_cv	(6) BL_pl	(7) Hand-picked
dair31_07	0.005 (0.029)		-0.014 (0.032)				
dair32_03	0.041** (0.017)	0.056*** (0.019)	0.038** (0.018)	0.056*** (0.019)	0.059*** (0.022)		0.049*** (0.016)
dair32_06	-0.049 (0.030)	-0.051* (0.027)		-0.048* (0.025)	-0.049* (0.026)		
dair31_05		0.002 (0.026)		0.007 (0.027)	0.002 (0.029)		
dpre32_02		-0.033*** (0.011)		-0.034*** (0.012)	-0.035*** (0.012)		
dpre32_12		0.001 (0.007)		-0.002 (0.009)	-0.003 (0.009)		
dpre32_11			0.001 (0.005)				
dpre31_04				-0.003 (0.006)	-0.004 (0.007)		
dpre32_01				-0.003 (0.010)	-0.004 (0.010)		
dair32_02					-0.014 (0.038)		
dair31_08						0.017 (0.041)	0.031 (0.037)
dair32_05						-0.032 (0.023)	
dpre31_01						0.001 (0.007)	
dpre31_08						0.007** (0.003)	
dpre32_05						-0.002 (0.005)	
dpre32_09						-0.009 (0.006)	
dair31_03							-0.009 (0.018)
dair32_07							-0.038 (0.025)
dpre31_07							-0.004 (0.004)
dpre32_07							0.001 (0.007)
N	6094	6094	6094	6094	6094	6094	6094
F-stat	7.94	24.17	5.16	26.4	26.44	12.00	15.07
Lasso controls	base	base	econ	econ	econ	none	none

Notes: The table reports results for considered Lasso specifications that pick fewer than 20 instruments. Base Lasso controls: climate zone indicator and level of grain procurement plan in 1932-33. Economic Lasso controls: percent drop in horse population in June 1928-June 1932 and indicator that district center has a railroad station or port. Instruments in columns 1-5 are picked using district-level data, and in column 6 are picked using village-level data. Instruments in column 7 are historically-motivated hand-picked instruments considered by Rozenas and Zhukov (2019). The sample consists of units with fewer than 20,000 residents in 2001. Conely standard errors adjusted for spatial correlation within radius of 50 km are in parentheses:* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Balance table: Blacklisting and pre-famine district-level characteristics

Variable	Blacklisted		$\widehat{Blacklisted}$		N
	Coef	S.e.	Coef	S.e.	
Population (1927)	-7958.6	(4781.8)*	-284404	(106562.1)***	6094
Rural population density (1930)	0.388	(1.543)	-60.429	(21.910)***	6094
Share of Ukrainians (1927)	-0.133	(1.238)	-78.151	(42.550)*	6094
Share of Russians (1927)	0.359	(0.675)	27.676	-20.987	6094
Share of Jews (1927)	-0.043	(0.221)	-1.544	-3.407	6094
Share of Germans (1927)	0.234	(0.665)	15.145	-20.23	6094
Direct access to railroad or port (1933)	-0.018	(0.05)	-0.931	-0.705	6094
Livestock per capita (1925)	0.007	(0.019)	0.583	(0.298)*	6094
Equipment per capita (1925)	0.009	(0.004)**	0.199	(0.069)***	6094
Share of grain land (1925)	-0.005	(0.009)	0.043	-0.19	5728
Share of grain land (1930)	0.005	(0.007)	0.201	-0.166	5148
Log(grain procurement plan) (1930-31)	-0.011	(0.112)	-2.5	-2.217	5911
Fulfillment of procurement plan (1930-31)	0.001	(0.013)	-0.59	(0.209)***	5911

Notes: The table reports the coefficient δ and corresponding standard error from regression $Y_d^v = \alpha + \delta \times X_{v,d} + \theta_1 lon + \theta_2 lat + error_v$ where Y_d^v stands for the pre-famine district-level characteristics and $X_{v,d} = \{BL_{v,d}, \widehat{BL_{v,d}}\}$. The slope corresponds to the weighted average difference in outcomes of districts with and without blacklisted communities. The sample consists of units with fewer than 20,000 residents in 2001. Conley standard errors (50 km) are reported in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Contribution of controls to the effect of blacklisting on nightlight intensity

	(1)	(2)	(3)	(4)	(5)	(6)
Blacklisted	-1.132	-1.631	-1.631	-1.705	-1.992	-1.707
	(0.479)**	(0.526)***	(0.526)***	(0.548)***	(0.694)***	(0.627)***
Centroid longitude	0.030	0.032	0.032	0.032	0.027	0.025
	(0.007)***	(0.010)***	(0.010)***	(0.009)***	(0.009)***	(0.008)***
Centroid latitude	-0.002	0.016	0.016	0.019	0.029	0.029
	(0.011)	(0.032)	(0.032)	(0.033)	(0.034)	(0.031)
climate==Boreal		-0.071	-0.071	-0.065	-0.028	-0.035
		(0.121)	(0.121)	(0.126)	(0.143)	(0.130)
climate==Boreal Steppe		-0.178	-0.178	-0.196	-0.182	-0.161
		(0.069)***	(0.069)***	(0.071)***	(0.072)**	(0.067)**
climate==Maritime Steppe		0.071	0.071	0.076	0.088	0.072
		(0.073)	(0.073)	(0.074)	(0.081)	(0.075)
Rural population density (1930)		0.008	0.008	0.006	0.006	0.006
		(0.003)***	(0.003)***	(0.003)**	(0.003)**	(0.003)**
Livestock per capita (1925)				-0.252	-0.439	-0.347
				(0.148)*	(0.248)*	(0.237)
Equipment per capita (1925)					1.369	0.953
					(1.097)	(0.990)
Direct access to railroad or port (1933)						0.108
						(0.033)***
Constant	1.157	0.117	0.117	0.106	-0.196	-0.211
	(0.563)**	(1.359)	(1.359)	(1.407)	(1.492)	(1.374)
N	6094	6094	6094	6094	6094	6094
F-stat	42.28	29.24	29.24	27.71	27.41	26.4

Notes: The table reports IV results for equations (1) and (2). The sample consists of units with fewer than 20,000 residents in 2001. Conley standard errors (50 km) are reported in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Effect of blacklisting on population size

	Log population		
	1959	1989	2001
Blacklisted	0.005	0.151	0.047
	(0.520)	(0.603)	(0.670)
N	5673	5673	5673
Mean (BL=0)	7.6	7.2	7.0
SD (BL=0)	.6	.6	.8
Mean (BL=1)	7.8	7.4	7.2
SD (BL=1)	.6	.8	.8

Notes: Table reports the IV estimate of β_1 from equation (1.1). Conley standard errors (50 km) are reported in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. First-stage F-statistic for excluded instruments is 42.27. Controls: Rural population density, livestock and equipment per capita, access to railroad or port, climate zone indicator, and centroid longitude and latitude. The regression is performed for villages with fewer than 20,000 residents in 2001 for which information on population is available in 1959, 1989, and 2001. The data from the 1989 and 2001 censuses are available from the Ukrainian State Statistics Committee. I created the village-level population count in 1959 by digitizing the archival documents. The information about the village-level population in 1959 remapped to Ukraine's administrative division in August 1988 is available from the Central State Archive of Higher-Level State Authorities in Kyiv, Ukraine (2184 pages). I manually matched the 1959 population count records to the administrative division of Ukraine in 1989 and, consequently, in 2017 (accounting for changes in the names of some units). Finally, I aggregated the population count in the villages to the village council level. While performing the matching, I was able to account for almost all villages that disappeared between 1959 and 2017 (583 villages on the territory of Soviet Ukraine in 1933). It mitigates the concern about measurement errors in historical population count.

Table A.5: Effect of blacklisting on change in district-level ethnic composition

	(1) Ukr	(2) Rus	(3) Jew	(4) Pol	(5) Ger	(6) Oth
Blacklisted	34.03	-7.95	1.53	-18.18**	-11.05	1.63
	(24.14)	(23.50)	(4.22)	(8.10)	(16.98)	(11.64)
N	5673	5673	5673	5673	5673	5673
Mean (BL=0)	4.4	3.4	-3	-1.8	-1.8	-1
SD (BL=0)	12.2	9.2	3	3.6	5.4	4.6
Mean (BL=1)	8.6	3.2	-3.2	-3	-4	-1.8
SD (BL=0)	14	8	2.8	6	7.8	5.8

Notes: Table reports the IV estimate of β_1 from equation (1.1). Conley standard errors (50 km) are reported in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. First-stage F-statistic for excluded instruments is 42.27. Controls: Rural population density, livestock and equipment per capita, access to railroad or port, climate zone indicator, and centroid longitude and latitude. The regression is performed for villages with fewer than 20,000 residents in 2001 for which information on population is available in 1959, 1989, and 2001. Change in ethnic composition is computed as a difference in share of a specific nationality in 2001 and 1927. First, I map the number of district residents with different nationalities in 1927 to present-day administrative division using spatial reweighting and aggregate the village-level number of residents with different nationalities in 2001 on the district level. Then, given the reweighted district-level population counts by nationalities in 1927 and 2001, I compute the shares of population with different nationalities and their change over time.

A.2 More Information on Weather and Blacklisting

A.2.1 Weather Patterns in Ukraine in the 1930s

Based on the data about monthly temperature and precipitation in Kyiv region, Davies and Wheatcroft (2016) conclude that Ukraine has experienced unfavorable weather (late spring and hot summer with unfavorable precipitation pattern) which negatively affected harvest for two years in a row, 1931-32. This appendix corroborates this conclusion by providing additional information about the weather patterns in the 1930s. It also demonstrates that my weather instruments (deviations from median of air temperature in March, May, and June and precipitation in January, February, April and December) capture weather aberrations rather than systematic fluctuations in weather.

Table A.6 summarizes median weather in 1926-1930 (considered as benchmark in the main analysis), weather volatility in 1901-1930, and deviations of weather in 1931-1932 from the benchmark. It allows to understand how weather in 1931-32 was different from the usual weather conditions.

Median Weather in 1926-1930. According to Panel A, the coldest month in Ukraine is typically February, followed by January, December and March. The average temperature in March is typically below zero. The hottest month is July followed by August and June. June is also characterized by most precipitation followed by other summer months and May. In winter months, December is characterized by the highest precipitation. Figure A.7 shows that these median values well represent monthly temperature and precipitation norms.

Deviations From Median in 1931-32. Panels B and C summarize the weather aberrations in 1931-32 relative to the median. In 1931, only in March the weather was close to median (although with more precipitation). May and June were warmer and dryer than median, and July was warmer than median but with more precipitation. All other months were colder than median and winter months were characterized by less precipitation. 1932 was characterized by long winter with February, March and April being colder and wetter than median. Overall, the first half of the year until July was characterized by more precipitation than usual. May, June and July were again warmer usual. The were followed by fall and winter warmer than median (except for November) and dryer than median (except for October).

Volatility. Panel D shows that air temperature in the most volatile in February and November, and precipitation is the most volatile in July and August.

Optimal Instruments. Lasso-selected instruments do not correspond to months when weather is typically most volatile. However, these are months when weather matters for agricultural operations. For example, March and April weather affects the timing of spring sowing, and thus, harvest from spring crops gathered in fall, and weather in winter matters for the germination of winter crops.

Table A.6: Monthly temperature and precipitation in 1926-1932 (district level)

Variable	Temperature, degree Celcius				Precipitation, mm			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Panel A. Median 1926-1930								
Jan	-6.0	1.2	-9.2	-2.3	2.3	0.5	0.6	3.4
Feb	-6.9	1.3	-10.1	-4.3	2.1	0.7	0.8	3.9
Mar	-1.4	1.1	-3.6	2.0	2.3	0.6	0.5	3.7
Apr	8.3	1.0	6.0	10.6	4.0	0.9	2.2	5.8
May	14.9	0.7	13.4	16.6	6.0	1.3	1.8	9.6
Jun	17.2	1.0	15.1	20.5	6.6	1.5	1.7	9.7
Jul	20.2	1.2	18.0	23.2	5.6	2.2	1.7	9.7
Aug	20.0	1.4	17.5	22.9	5.4	1.3	2.6	8.6
Sep	13.6	1.3	11.1	17.8	3.8	1.1	0.9	6.5
Oct	9.2	1.0	7.1	12.6	3.7	0.8	2.1	5.5
Nov	4.3	1.1	1.7	8.3	3.9	0.7	2.3	5.8
Dec	-3.3	1.6	-7.8	0.7	4.5	0.7	2.7	6.9
Panel B. Deviation from median, 1931								
Jan	-0.5	0.7	-1.5	1.2	2.3	1.5	-1.5	8.1
Feb	-1.8	1.4	-4.6	1.9	-0.9	0.6	-2.6	1.2
Mar	0.1	1.2	-1.5	3.6	1.7	1.0	-0.2	5.0
Apr	-3.0	0.5	-4.1	-1.6	0.7	1.2	-2.2	4.8
May	1.9	0.6	0.4	3.1	-0.9	1.8	-5.3	3.1
Jun	1.6	0.5	0.2	3.0	-1.2	2.0	-5.3	8.9
Jul	2.2	0.7	1.1	3.8	1.6	2.6	-3.4	9.1
Aug	-1.1	0.3	-1.9	-0.4	1.3	2.2	-3.5	8.3
Sep	-0.7	1.1	-2.9	1.8	2.1	1.6	-1.5	5.9
Oct	-1.8	0.4	-2.3	-0.5	0.1	0.7	-2.1	3.1
Nov	-4.4	0.2	-4.9	-3.8	-0.1	1.9	-3.5	6.2
Dec	-1.0	1.0	-3.0	1.3	-1.4	1.1	-3.4	2.0
Panel C. Deviation from median, 1932								
Jan	3.2	1.2	1.2	5.7	0.4	0.9	-2.1	2.7
Feb	-3.6	0.5	-5.1	-2.0	0.7	0.7	-0.9	3.4
Mar	-3.0	0.8	-5.0	-1.3	1.0	1.0	-1.8	4.0
Apr	-0.7	0.5	-2.0	0.8	2.6	2.6	-3.2	7.3
May	1.3	0.7	-0.4	2.4	1.6	1.9	-2.8	8.0
Jun	*1.2*	1.3	-1.0	4.1	6.4	2.9	-0.6	15.4
Jul	0.4	0.8	-1.0	2.0	1.0	1.7	-2.6	9.6
Aug	-0.4	0.5	-1.4	0.8	-0.3	1.8	-3.9	6.1
Sep	3.0	0.5	1.7	3.9	-1.0	1.0	-4.3	2.2
Oct	1.4	0.5	0.6	2.8	2.3	2.4	-1.8	8.0
Nov	-2.3	0.3	-3.0	-1.6	-0.2	1.5	-2.8	2.6
Dec	2.6	0.8	0.2	4.5	-1.4	1.0	-3.9	1.9
Panel D. Standard deviation, 1901-1930								
Jan	2.7	0.2	2.2	2.9	1.7	0.3	1.2	3.0
Feb	3.8	0.2	3.4	4.4	1.6	0.2	1.2	2.6
Mar	2.6	0.2	2.2	3.0	1.6	0.3	1.1	2.7
Apr	1.9	0.3	1.3	2.4	1.9	0.3	1.1	3.1
May	2.0	0.3	1.3	2.5	2.4	0.4	1.2	4.0
Jun	1.6	0.2	1.1	2.0	2.6	0.4	1.8	4.0
Jul	1.1	0.1	0.9	1.3	3.3	0.5	2.3	4.8
Aug	1.6	0.2	1.2	2.0	2.6	0.5	1.7	4.7
Sep	1.8	0.1	1.6	2.0	2.9	0.4	2.0	4.1
Oct	2.3	0.1	2.0	2.6	2.3	0.3	1.4	3.2
Nov	2.9	0.1	2.7	3.1	2.3	0.3	1.6	3.4
Dec	2.4	0.1	2.1	2.7	1.9	0.2	1.4	2.7

Notes: The table is calculated by the author using the gridded weather data for 1901-1932 from Matsuura and Willmott (2014). The instruments are temperature deviations in May 1931, March and July 1932 and precipitation deviations in April 1931, January, February, and December 1932.

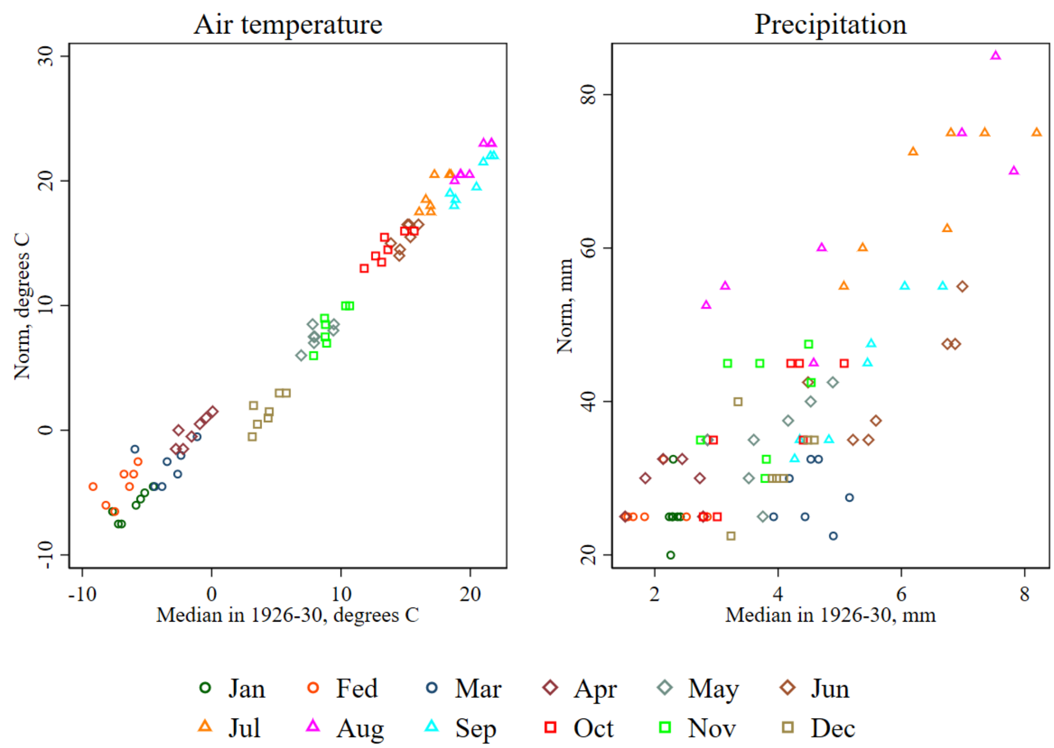


Figure A.7: Weather median versus norms (month-by-region)

Notes: The figure demonstrates that median weather in 1926-1930 well represents the weather norms in 8 regions of Ukraine in the early 1930s. Norm denotes the average values of temperature and precipitation considered normal in Soviet Ukraine in a given month. Constructed by the author using data from Asatkina (1935).

A.2.2 Weather Instruments

This appendix provides additional information about weather instruments.

Figure A.8 visualizes temperature deviations in two months that have the largest predictive power in the first stage – March and June 1932 (see Table 1.3). Consistent with conclusions of Davies and Wheatcroft (2016), the figure shows that March was colder than usual and June was warmer than usual on the most territory of Soviet Ukraine.

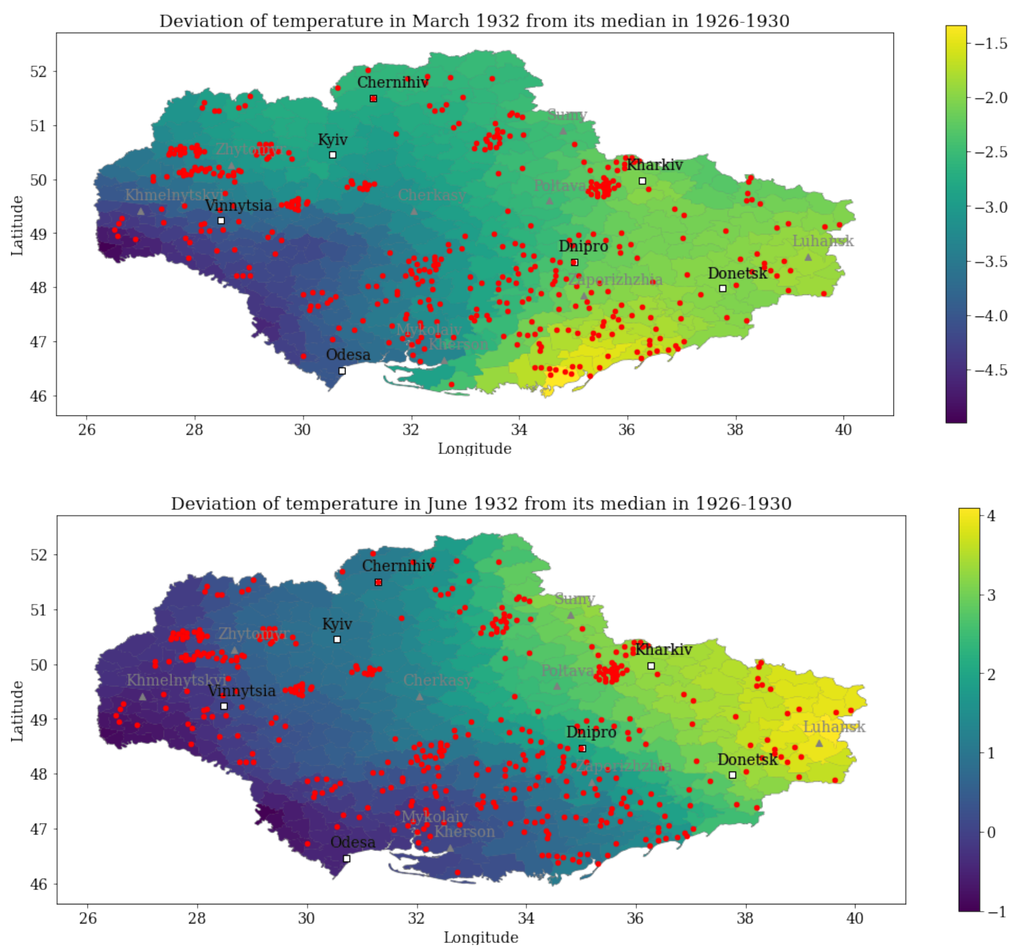


Figure A.8: Temperature deviations from the median in March and June 1932

Notes: Based on data from Matsuura and Willmott (2014). Red dots denote centroids of blacklisted village councils.

Figure A.9 visualizes the predicted values of blacklisting given the weather and other controls. Figure A.10 reveals positive spatial correlation between the actual and predicted blacklisting status. These results corroborate the conclusion that the preferred set of instruments produces a strong first stage.

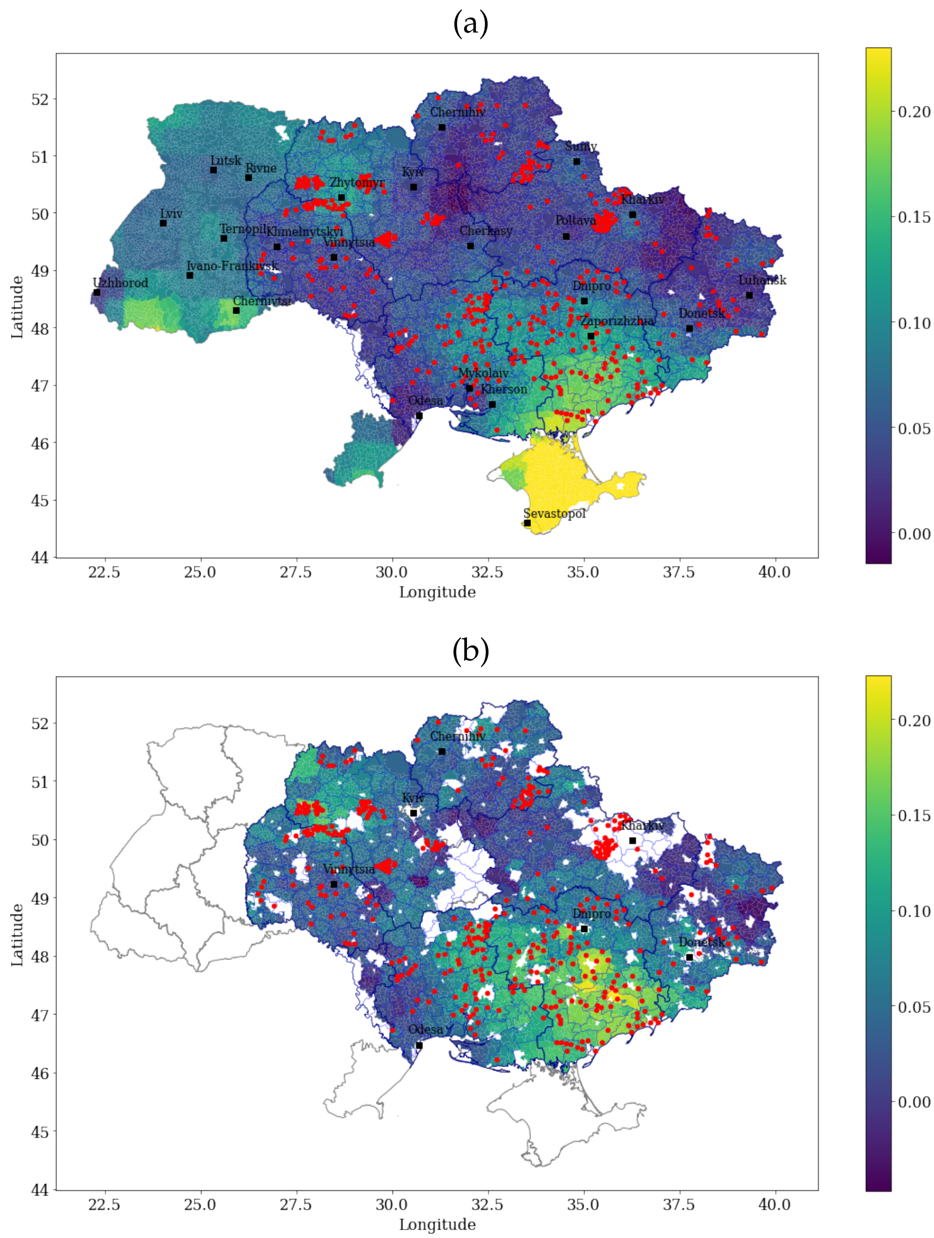


Figure A.9: Fitted value of blacklisting using weather-based instruments in 1931-32

Notes: Based on data from Matsuura and Willmott (2014). Red dots denote centroids of blacklisted village councils. Panel (a) reports blacklisting probability using information only about weather without other controls. Panel (b) reports the blacklisting probability using both weather instruments and controls for localities with fewer than 20,000 residents in 2001. The missing values are mostly due to missing information on historical controls (such as district-level equipment per capita and livestock per capita in 1925). .

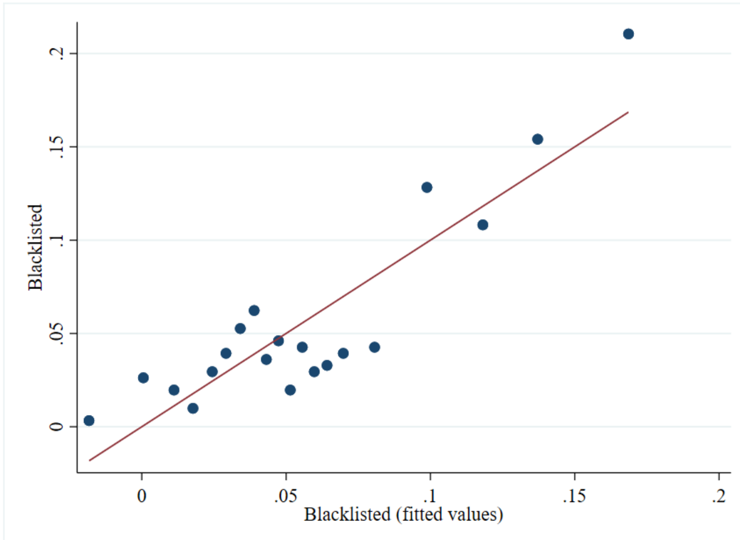


Figure A.10: Binscatter for actual and fitted probability of blacklisting

Notes: The binscatter shows the association between the probability blacklisting predicted using all weather instruments in 1931-32 and controls for localities with fewer than 20,000 residents in 2001 and actual blacklisting status.

A.2.3 Robustness of the First Stage to Lagged and Leading Weather Shocks

Next, I examine the effect of weather shocks in 1931-32 versus preceding and following years on blacklisting to verify that the power of the first stage is driven by the 1931-32 shocks. In each specification, I include pseudo-instruments (deviations of weather from median in 1926-30 for different years) in addition to my instrument set and report joint F-test for my main instruments set and pseudo-instruments separately according to equation below:

$$BL_{v,1933} = \alpha + \sum_{m}^{Instr} \gamma_{m,1931-32} \cdot DevFromMedian_{m,1931-32} + \sum_{m}^{Instr} \delta_{m,y} \cdot DevFromMedian_{m,y} + \mathbf{X}_v'\theta_B + \nu_v$$
$$\text{where } y = \{1920-21, 1922-23, 1936-37, 1938-39\}$$

The results are reported in Table A.7. It shows that even after controlling for weather in the same months as selected by Lasso but in different years, the true instruments have most predictive power for blacklisting status. Moreover, controlling for additional weather shocks does not change the coefficients on the true instruments qualitatively: colder than usual March and warmer than usual June still decrease the probability of blacklisting just as more precipitation in February does.

Table A.7: Robustness of the first stage results to shocks in different years

	Depvar: *Blacklisted*					
	(1)	(2)	(3)	(4)	(5)	(6)
Deviation of temperature from median						
May'31	0.018	-0.009	-0.049	-0.004	0.002	-0.026
	(0.023)	(0.024)	(0.035)	(0.042)	(0.037)	(0.060)
Mar'32	0.060***	0.046*	0.050*	0.052*	0.087***	0.029
	(0.018)	(0.027)	(0.030)	(0.029)	(0.025)	(0.044)
Jun'32	-0.049**	-0.059**	-0.040	-0.078***	-0.059*	-0.099**
	(0.024)	(0.024)	(0.029)	(0.029)	(0.030)	(0.048)
Deviation of precipitation from median						
Apr'31	-0.005	0.009	-0.000	0.002	-0.005	0.000
	(0.005)	(0.009)	(0.006)	(0.007)	(0.006)	(0.009)
Jan'32	-0.003	0.002	-0.008	0.015	-0.000	0.017
	(0.009)	(0.010)	(0.009)	(0.015)	(0.009)	(0.018)
Feb'32	-0.035***	-0.053***	-0.037***	-0.041***	-0.026*	-0.049***
	(0.012)	(0.013)	(0.009)	(0.016)	(0.013)	(0.014)
Dec'32	-0.001	0.007	0.001	0.006	0.006	-0.023
	(0.008)	(0.011)	(0.012)	(0.012)	(0.009)	(0.017)
N	6094	6094	6094	6094	6094	6094
F-stat'31-32	42.00	35.42	29.71	34.79	51.04	17.03
F-stat'20-21		11.00				11.00
F-stat'22-23			11.11			11.11
F-stat'35-36				10.83		10.83
F-stat'37-38					14.67	14.67

Notes: The sample consists of units with fewer than 20,000 residents in 2001. Conley standard errors (50 km) are parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard controls are included in all specifications. In addition, columns (2)-(5) include controls for weather deviations in one of following intervals {1920-21, 1922-23, 1936-37, 1938-39} (i.e., there 14 weather deviations per specification in total). Column (6) controls for all weather deviations included in columns (2)-(5) at the same time (i.e., 70 weather deviations in total).

A.3 Elasticity of Real Gross Regional Product With Respect to Nightlight Intensity

This appendix discusses my approach to calculate the elasticity of real gross regional product with respect to nightlight intensity to facilitate economic interpretation of the main regression results.

Economists estimate the conversion rate between nightlight intensity and economic activity using linear regression either on national or sub-national level (e.g. Henderson et al., 2012; Bluhm and McCord, 2022). Literature points a number of issues for estimating this relationship correctly. First, different satellites may have different calibrations so one unit of nightlight intensity may correspond to different amounts of lights emitted by the source. For this reason, researchers control for the satellite fixed effects. Second, the relationship between nightlights and GDP may depend on the economy structure. For example, nightlights are less responsive to changes in GDP in more agricultural areas (Bluhm and McCord, 2022; Ghosh et al., 2010). As structure of economic activity changes, the relationship between nightlights and GDP may change over as well. A way to address it is to control for observable local economic conditions, such as share of agriculture and/or time and location fixed effects.

I estimate the elasticity of real gross regional product in a region r to nightlights recorded by a satellite s at time t using the following regression equation:

$$ln(realGRP_{rt}) = \alpha_r + \tau_t + \zeta_s + \beta log(NL_{srt}/area_r) + \gamma_1 shagr_{rt} + \varepsilon_{rt} \tag{A.1}$$

where $log(NL/area)$ is the nightlights per well-lit pixel each having area of roughly one sq. km. and $shagr$ is the share of agriculture in GRP. I perform regression for a panel of 25 regions of Ukraine in 2001-2013. The unit of observation is region by satellite-year. Inclusion in my analysis of the data recorded by multiple satellites increases my overall sample size to 500 observations. I obtain the real gross regional product series in 2001 prices by adjusting the gross regional product in each year by GDP deflator. The data about the GDP, deflator, and share of agriculture are obtained from the State Statistics Service of Ukraine data. Figure A.11 shows that there is positive relationship between regional nightlights and real gross regional product.

According to the regression results summarized in Table A.8, the elasticity of real gross regional product to brightness of nightlights per well-lit pixel in Ukraine is between 0.13 and 0.31. Inclusion of the satellite fixed effects increases the estimated elasticity whereas controlling for the year and region fixed effect reduces the estimated elasticity [1] My results are consistent with the estimates in prior literature. Henderson et al. (2012) performing similar analysis for a panel of 188 countries in 1992-2008 finds elasticity estimate about 0.28. The elasticity is even higher in the subsample of low and middle income countries (about 0.31). These estimates are are often used as a benchmark in the literature. Similar

[1] The images for 2001-2013 were taken by one of four satellites (F14, F15, F16, F17).

approach applied to the GRP of Indian states in 1992-2013 yields an estimate of elasticity between 0.15 and 0.18 (Prakash et al., 2019). For Dominican Republic, Ishizawa et al. (2019) finds the elasticity of quarterly aggregate GDP per area with respect to average nightlights equal to 0.11.

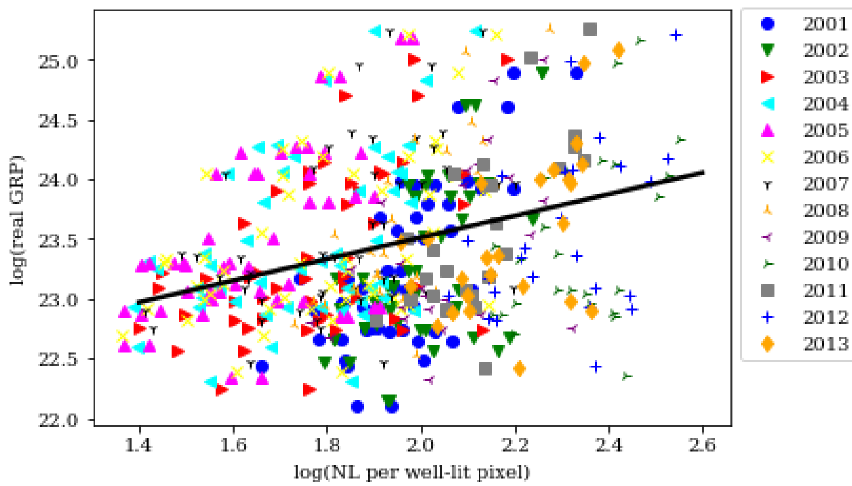


Figure A.11: Real gross regional product and nightlight intensity per well-lit pixel in Ukraine in 2001-2011

Notes: Based on region-level data from State Statistics Service of Ukraine. The linear fit line corresponds to estimates in column 1 of Table A.8.

Table A.8: Elasticity of real gross regional product to nightlight intensity per well-lit pixel in Ukraine in 2001-2013

	log(real gross regional product)				
	(1)	(2)	(3)	(4)	(5)
log(NL/pixels)	0.901***	0.731***	0.150**	0.309***	0.126***
	(0.112)	(0.090)	(0.059)	(0.082)	(0.036)
Area		0.061***	0.039***	0.038***	0.549***
		(0.003)	(0.002)	(0.002)	(0.025)
% agriculture			-0.059***	-0.060***	-0.007***
			(0.002)	(0.002)	(0.001)
N	500	500	500	500	500
R-sq	0.11	0.42	0.79	0.79	0.99
Satellite FE	No	No	No	Yes	Yes
Year FE	No	No	No	No	Yes
Region FE	No	No	No	No	Yes

Notes: Robust standard errors are in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is a panel of 25 regions of Ukraine in 2001-2013. The unit of observation is region by satellite-year.

I will use the elasticity estimates 0.13 (from column 5 of Table A.8) and 0.31 (Henderson et al., 2012, Table 4, column 1) for interpretation of economic meaning of my main regression result.

Appendix B

Appendix for Chapter 2

B.1 Examples of Survey Questions

Although the exact formulation of survey questions may vary across survey waves and countries, the surveys are generally comparable in terms of question wording. Listed below is a typical wording of questions according to wave 12. Detailed information about country-specific survey details can be found in the codebooks here: https://www.gesis.org/en/issp/modules/issp-modules-by-topic/family-and-changing-gender-roles.

The survey questions about **time allocation** are:

- Respondent:

  [Market work] *How many hours, on average, do you usually work for pay in a normal week, including overtime? If you work for more than one employer, or if you are both employed and self-employed, please count the total number of working hours that you do.*

  [Household work] *On average, how many hours a week do you personally spend on household work, not including childcare and leisure time activities?*

  [Family care] *On average, how many hours a week do you spend looking after family members (e.g. children, elderly, ill or disabled family members)?*

- Spouse:

  [Market work] *How many hours, on average, does your spouse/ partner usually work for pay in a normal week, including overtime? If he/ she works for more than one employer, or if he/ she is both employed and self-employed, please count the total number of working hours that he/ she does.*

  [Household work] *And what about your spouse/ partner? On average, how many hours a week does he/ she spend on household work, not including childcare and leisure time activities?*

[Family care] *And on average, how many hours a week does he/ she spend looking after family members (e.g. children, elderly, ill or disabled family members)?*

The survey question about **income** for Germany (the formulation varies by country; in particular, surveys in some countries as about net income while in others – about gross income):

[Household income] *How high is the total net monthly income of your household? By this I mean the amount remaining after deductions for tax and social security contributions. For self-employed, please ask for average net monthly income, after deductions for overheads.*

[Respondent's income] *How high is your own net monthly income? By this I mean the amount remaining after deductions for tax and social security contributions. For self-employed, please ask for average net monthly income, after deductions for overheads.*

The survey questions about **gender attitudes**:

- And to what extent do you agree or disagree...?

  [Both should work] *Both the man and woman should contribute to the household income.*

  [Traditional division] *Men's job is to earn money; women's job is to look after the home and family.*

- Options: 1 – Strongly agree, 2 – Agree, 3 – Neither agree nor disagree, 4 – Disagree, 5 – Strongly disagree, 8 – Can't choose, 9 – No answer.

B.2 Data Processing

Imputation Procedures for Hours

The information on hours devoted to work for pay is quite noisy and sometimes contradicts answers to other questions (e.g., employment status or income). In the raw data, there is a substantial number of respondents with missing information on hours in 2002. Altogether, this results in a substantial difference in the level of hours worked in 2002 and 2012. For this reason, I employ several cross-checks and imputation procedures in the data. I report summary statistics for resulting hours worked by country in Appendix Table B.2.

I refer to the raw unimputed data as "original data". Panel A of Appendix Table B.2 shows that hours worked for pay by women are smaller on average by up to 15 hours. This is mainly due to missing information on hours worked for many respondents. To mitigate the impact of missing observations, I utilize information on employment status.

First, if a person reports employment status out of the labor force or unemployment, I set the number of hours worked for pay to zero. This sets all missing and non-missing positive values to zero. I make an exception for students, trainees, and apprentices because they could combine their studies with work. If students report working nonzero hours for pay, I do not change this value. If students do not report hours worked, I assume this means they do not work for pay and, therefore, replace missing values with zeros. I refer to the resulting hours series as "lower bound" and summarize them in panel B of Appendix Table B.2. Now the average hours worked for pay by women are similar in 2002 and 2012 and more observations are retained in the sample. When computing the "lower bound" hours series, I did not apply any imputation procedure for hours for people who report being employed for pay but do not report how many hours they work. To construct the "upper bound" hours series, for all respondents who report working for pay and for whom hours are missing or zero, I assign a value of 40 hours. If information on hours worked for pay is not missing for employed individuals, I keep it as it is. Summary statistics for the resulting hours series is reported in panel C of Appendix Table B.2. The "upper bound" hours are very similar to the "lower bound". Given that the "upper bound" series was subject to imputation of hours for employed, unemployed, and out of the labor force, this series is used throughout the paper for both hours worked by respondents and their spouses.

For hours devoted to household work and family care, there are no variables that would allow for cross-validating the values, so the paper relies on the data as reported by respondents with the exception that missing values are interpreted as zero unless they correspond to a category denoted in the codebook as respondent's refusal or unwillingness to answer the question.

Computing Female Relative Time Input to Different Activities

For female respondents, female time share in the household total is the amount of time that a respondent devotes to a particular activity divided by the sum of time devoted to the activity by respondent and spouse. For male respondents, female time share in household total input is the amount of time that a spouse devotes to a particular activity divided by the sum of time devoted to the activity by respondent and spouse.

Computing Female Income Share

I assume that the income of household members other than husband and wife is negligible. The fact that the average household size is close to 3 and couples have on average one child (Appendix Table B.1) supports this assumption. Therefore, for female respondents, I compute female income share as a ratio of respondent's income to household income. For male respondents, I compute female income share as family income net of respondent's income divided by family income.

B.3 Notes About Construction of Variables

Gender. There are a few observations with missing gender which I omit throughout the paper.

Children. The information about the number of children by age 0-6 and 6-18 years old is available only in waves 2002 and 2012. For 1994, only information on whether a person has ever had children is available which is not very informative.

- In 2002, missing values in the number of children were not clearly coded and may stand for one of the following: "not applicable", "don't know", "refuse to answer". Taking into account that in 2012 data respondents are much more likely to have no children than to refuse to answer this question, I impute zero children if the answer is missing. The number of children is not available for Bulgaria in 2002.
- In 2012, if a person has no children, this is coded as 0. In the sample, about 2/3 of respondents reported having 0 children, whereas less than 3% of respondents refused to answer the question about the number of children. In some cases, it is possible to impute the remaining missing values given the information on household size. In 2012 (and in Russia in 2002), household size corresponds to the number of people residing in the household including children. If in the household of a married respondent currently residing with a spouse the household size is 2, this suggests that they did not have children at the moment of the survey.

Religion. I regrouped religion into five categories – no religion, Roman Catholic, Protestant, Christian Orthodox, and other religions. The latter category includes Muslim which is the dominant religion in Israel. Information about religion is not available for Slovenia in 2002.

The mother's work status during respondent's childhood is summarized by a dummy variable that takes a value of 1 if the respondent says their mother worked for pay before respondent turned 14. In all other cases, it takes a value of zero. This includes a few (about 1% or less) occasions when a respondent reports that their mother was not present during their childhood which is coded in the same way as if the mother was not employed for pay. Additionally, in 2012, about 4.5% of respondents answered "don't know" or refused to answer the question about mother's employment which is also coded in the same way as it is when mother was not employed for pay.

Education. Information about the respondent's education is available from a homogenized variable from all 3 waves. However, the homogenized variable does not indicate any person from Russia, Israel, or Great Britain as having higher education in the 2012 wave (there is no such issue for other countries or previous years). I infer this information for these three countries from country-specific education variables instead (not available for the previous waves). Information about the spouse's education is available in 2002 but very limited in 2012 and, therefore, not used in the analysis.

Employment status. The information about respondent's and spouse's employment status allows me to classify respondents and spouses into employed, unemployed, and out of labor force. This variable is used to cross-validate hours worked for pay.

Hours worked for pay. The information about hours worked by respondents is missing for BG, DE, and ES in 1994 and available for all countries in 2002 and 2012. Hours worked by spouse are missing for CZ, SI, SK in 2002 and for GB and BG in 2012). For this reason, they are not included in the regression sample.

Hours spent on household chores (other than family care) are available in 2002-2012 for respondents and their spouses from all the countries.

Hours spent on family care are available for both respondents and spouses only in 2012.

Respondent's earnings in 1994 are available for all countries but ES, in 2002 it is missing for NL, and available for all countries in 2012. However, there are a lot of missing values.

a. Information on income is missing for 4,993 out of 19,981 respondents of non-retirement age in the 2002 sample. In Britain, information about earnings is available only for people who are employed, so I substitute missing earnings of non-employed respondents to 0. This imputation procedure reduces the number of missing observations to 4695.

b. In 2012, 2,953/17,491 observations for respondents of non-retirement age are missing. Unlike missing observations in 2002, in 2012 they are marked by flags as "don't know" or "refuse to answer". Therefore, no imputation is appropriate in this case.

Family income is missing for IL in 1994 and available for all countries in 2002 and 2012. In all countries in the sample except for Russia, income of all members is reported; in Russia in 1994 – 2002, income is reported per household member; therefore, I multiply the reported variable by household size to obtain the household income. In most countries, the question refers to net household income after taxes and deductions (AT, BG, CZ, DE, LV, PL), although in some countries income before taxes or deductions is reported (GB, NO, SE, US). It is not clear from the documentation whether income data reported is gross or net income in HU and RU.

- Information on family income is missing for 3,482/19,981 respondents in 2002. In 2012, missing income is clearly coded as "don't know", "refuse", etc. so I also keep it as missing without any imputations. This leaves me with 3,916/17,491 missing observations.

Income share. In some countries, the currency could have changed over time, so it is more reliable to focus on share rather than level. Within a given year, both respondent's income and family income are reported consistently in the same currency in a given country.

- In some countries, income share exceeds 1, which is implausible. This accounts for about 1.5% of observations for a sample of non-retirement age respondents in 2002-2012. Therefore, I exclude such observations from the analysis. Altogether, this leaves me with about 15,201 married respondents of non-retirement age whose income share does not exceed 1.

- In the three-factor framework, it is necessary to distinguish between female and male input to household income. The data unambiguously provides information about respondents income and total family income. To maximize sample size, I create an imputed income share variable where I assume that family income consists of only respondent's and spouse's income and, therefore impute spouse's income share as 1 – respondent's income share (for respondents with income share less than 1). If the respondent is female, I impute the income share of the male spouse from the household, and vice versa. The imputation reduces the resulting income share slightly but provides plausible values for the average income shares.

Gender attitudes are available for all countries and years and are coded as categorical variables defining a degree of agreement with a particular statement.

Who does chores? A person in the household responsible for doing a particular task (cooking, cleaning, laundry, etc.) is available for all countries in 1994-2012. There are a few missing values in 2002.

Urban-rural status of the settlement comes from a variable defined in the survey. In different countries, the definition of city and village may correspond to different population size cutoffs. The data do not allow me to classify the settlements using a homogenous cutoff across all countries (e.g., settlements with 5000 citizens or less and settlements with at least 100 000 citizens). Information on urban-rural status is missing for Germany, Poland, Russia, and Israel in 2002. For this purpose, I use the information about the respondent's settlement population in these four countries to classify the settlement into urban, suburban/town, or rural areas. This variable is meant to capture differences in lifestyle resulting from living in urban versus rural areas. For example, people in rural areas may have a subsidiary plot, and therefore time allocation to household work, although unpaid, may improve household wellbeing via provision of food.

B.4 Tables and Figures

Table B.1: Country-level summary statistics in 2002-2012

	Total sample				Married sample		Married, non-retirement age					
Country	Sample size	Fem., %	Married, %	Age, mean	Fem., %	Nonretirement age, %	Sample size	Below secondary educ	Complete college, %	Household size, mean	Children, mean	Urban, %
2002												
DE-E	431	0.49	0.61	48.33	0.51	0.6	157	0.64	0.22	3.29	0.85	0.2
HU	1023	0.59	0.55	49.55	0.52	0.65	363	0.47	0.06	3.66	0.99	0.23
BG	1003	0.57	0.69	49.58	0.55	0.66	459	0.22	0.18	3.71	.	0.12
LV	1000	0.58	0.53	42.68	0.51	0.79	419	0.1	0.22	3.59	1.15	0.41
PL	1252	0.58	0.6	47.84	0.54	0.76	570	0.48	0.11	4.16	1.32	0.28
RU	1798	0.61	0.52	46.88	0.54	0.72	674	0.06	0.24	3.48	1.02	0.45
SI	1090	0.54	0.61	46.49	0.51	0.67	438	0.5	0.13	3.89	0.97	0.37
SK	1133	0.52	0.6	43.06	0.54	0.78	527	0.49	0.13	3.73	1.18	0.17
CZ	1289	0.63	0.58	42.92	0.64	0.79	591	0.53	0.11	3.42	0.99	0.23
US	1171	0.58	0.48	44.94	0.58	0.75	419	0.08	0.25	3.14	0.95	0.48
GB	1960	0.57	0.55	48.69	0.54	0.68	743	0.46	0.21	3.25	1.07	0.07
DK	1379	0.55	0.55	46.51	0.54	0.63	468	0.07	0.15	3.44	1.23	0.16
FI	1353	0.55	0.68	44.2	0.55	0.73	660	0.15	0.18	3.11	0.95	0.63
NO	1475	0.53	0.57	45.39	0.52	0.69	577	0.22	0.3	3.59	1.33	0.21
SE	1080	0.54	0.74	47	0.53	0.7	548	0.44	0.28	3.29	1.11	0.32
FR	1903	0.66	0.59	44.77	0.61	0.71	757	0.4	0.27	3.66	1.35	0.32
DE-W	936	0.52	0.59	46.42	0.52	0.64	358	0.7	0.15	3.44	1.12	0.23
AT	2047	0.62	0.5	45.91	0.59	0.73	749	0.72	0.09	3.58	1.1	0.17
ES	2471	0.52	0.56	45.99	0.52	0.65	892	0.61	0.1	3.65	0.96	0.33
IL	1205	0.56	0.66	42.41	0.57	0.75	593	0.27	0.3	4.4	1.89	0.43
PT	1092	0.59	0.58	47.68	0.57	0.65	415	0.67	0.14	3.54	0.95	0.28
2012												
DE-E	558	0.53	0.54	50.92	0.5	0.47	140	0.01	0.19	3	0.71	0.17
HU	1012	0.52	0.42	48.09	0.45	0.66	280	0.5	0.05	3.59	1.17	0.32
BG	1003	0.58	0.62	51.93	0.55	0.62	388	0.18	0.26	3.41	0.89	0.52
LV	1000	0.58	0.47	44.12	0.56	0.74	352	0.09	0.28	3.28	0.91	0.41
PL	1115	0.54	0.59	47.8	0.53	0.63	411	0.05	0.24	4.09	1.19	0.24
RU	1525	0.64	0.43	47.98	0.58	0.73	479	0.04	0.33	3.21	0.85	0.5
SI	1034	0.54	0.65	51.04	0.5	0.61	407	0.35	0.02	3.64	0.91	0.14
SK	1128	0.54	0.64	51.93	0.47	0.64	456	0.4	0.13	3.88	0.94	0.08
CZ	1804	0.55	0.57	47.42	0.52	0.67	679	0.3	0.09	3.34	0.9	0.37
US	1302	0.54	0.45	47.63	0.54	0.68	394	0.12	0.13	3.2	1.01	0.29
GB	950	0.54	0.47	52.57	0.45	0.49	221	0.26	0.34	3.43	1.16	0.07
DK	1403	0.51	0.51	46.2	0.53	0.61	438	0.08	0.19	3.5	1.71	0.16
FI	1171	0.56	0.54	47.1	0.54	0.54	334	0.05	0.2	3.54	1.33	0.05
NO	1444	0.52	0.59	47.97	0.48	0.57	474	0.18	0.46	3.72	1.34	0.19
SE	1059	0.54	0.54	52.01	0.5	0.51	286	0.29	0.36	3.76	1.15	0.26
FR	2409	0.65	0.57	51.84	0.62	0.56	756	0.31	0.26	4.08	1.34	0.13
DE-W	1208	0.51	0.56	49.21	0.5	0.59	399	0.09	0.24	3.43	1.06	0.17
AT	1182	0.55	0.61	47.99	0.52	0.68	486	0.69	0.14	3.04	0.76	0.38
ES	2595	0.53	0.61	49.17	0.51	0.63	1003	0.4	0.15	3.54	1.02	0.19
IL	1220	0.56	0.64	45.82	0.56	0.67	526	0.32	0.35	4.64	2.06	0.47
PT	1001	0.55	0.49	49.2	0.51	0.59	287	0.52	0.1	3.25	0.86	0.2

Notes: ISSP module "Changing family and gender roles". No survey weights are applied. The sample for which summary statistics are reported is provided in the column title.

Table B.2: Time devoted by married women to market work by country in 2002-2012 using different imputation procedures

Country	Year	Panel A: Original data				Panel B: Lower bound				Panel C: Upper bound			
		Abs., hours	Gap, hours	Rel., share	N	Abs., hours	Gap, hours	Rel., share	N	Abs., hours	Gap, hours	Rel., share	N
Post-socialist													
DE-E	2002	38.49	-6.86	0.46	96	28.89	-10.2	0.41	116	28.89	-10.2	0.41	116
	2012	28.32	-12.31	0.4	119	27.3	-12.99	0.39	122	28.05	-12.44	0.4	123
HU	2002	41.73	-7.27	0.46	134	26.58	-11.87	0.41	250	26.97	-11.55	0.41	273
	2012	26.23	-7.92	0.41	101	26.23	-7.92	0.41	101	26.4	-7.8	0.41	102
BG	2002	40.5	-2.78	0.48	179	29.14	-3.08	0.48	302	29.46	-2.92	0.48	315
	2012	28.35	.	.	144	28.35	.	.	144	28.51	.	.	146
LV	2002	40.9	-6.4	0.46	182	31.21	-12.47	0.39	259	31.64	-12.74	0.39	278
	2012	30.52	-7.83	0.42	212	30.52	-7.83	0.42	212	30.42	-8.12	0.42	227
PL	2002	42.8	-6.87	0.46	312	34.39	-9.19	0.44	327	34.49	-9.07	0.44	333
	2012	28.11	-12.04	0.4	223	28.11	-12.04	0.4	223	28.11	-12.04	0.4	223
RU	2002	40.29	-3.23	0.48	247	34.45	-5.67	0.45	311	34.48	-5.44	0.45	324
	2012	25.67	-14.87	0.37	268	25.67	-14.87	0.37	268	30.01	-10.4	0.42	308
SI	2002	40.83	.	.	94	35.87	.	.	107	35.87	.	.	107
	2012	29.61	-7.48	0.43	234	29.31	-7.85	0.42	235	29.36	-8.5	0.42	236
SK	2002	40.05	.	.	216	34.74	.	.	248	34.84	.	.	253
	2012	28.84	-7.69	0.43	235	28.6	-7.42	0.44	237	30.42	-5.98	0.45	255
CZ	2002	42.39	.	.	57	21.96	.	.	110	31.02	.	.	221
	2012	34.35	-9.38	0.41	402	34	-9.37	0.42	402	34.12	-9.14	0.42	422
Liberal													
US	2002	36.43	-10.22	0.43	218	29.11	-14.49	0.38	292	29.28	-14.33	0.38	297
	2012	25.11	-15.39	0.36	315	24.26	-16.01	0.35	317	24.53	-16.13	0.35	322
GB	2002	32.37	-13.47	0.4	234	24.05	-17.65	0.35	337	24.06	-17.68	0.35	341
	2012	20.48	.	.	84	20.48	.	.	84	20.48	.	.	84
Nordic													
DK	2002	35.94	-6.7	0.46	404	33.22	-7.28	0.45	408	33.52	-7.12	0.45	424
	2012	33.54	-4.52	0.47	417	32.25	-4.88	0.47	417	32.19	-5.09	0.46	429
FI	2002	36.38	-3.75	0.48	331	28.75	-7.61	0.43	435	28.9	-7.6	0.43	475
	2012	32.2	-6.89	0.45	286	32.01	-6.84	0.45	287	32.92	-5.93	0.46	290
NO	2002	33.09	-9.85	0.43	436	29.19	-12.34	0.4	452	29.12	-12.44	0.39	457
	2012	35.49	-7	0.44	400	33.03	-7.4	0.44	405	34.01	-6.9	0.44	420
SE	2002	36.04	-5.26	0.47	378	31.87	-8.05	0.43	394	31.94	-8.07	0.43	407
	2012	31.85	-8.84	0.42	231	31.57	-8.95	0.42	231	33.4	-7.6	0.43	239
Conservative													
FR	2002	33.78	-8.18	0.44	404	27.82	-11.63	0.4	513	28.31	-11.36	0.4	552
	2012	29.5	-10.28	0.41	533	28.72	-10.09	0.41	534	29.13	-9.74	0.42	541
DE-W	2002	30.98	-13.4	0.4	169	16.5	-25.17	0.25	218	16.42	-25.24	0.25	219
	2012	18.74	-21.94	0.29	323	17.99	-22.17	0.29	326	20.79	-19.77	0.32	326
AT	2002	32.26	-11.42	0.41	276	24.29	-14.05	0.38	378	24.64	-13.67	0.38	390
	2012	23.5	-13.13	0.38	282	23.5	-13.13	0.38	282	23.72	-12.96	0.38	286
Mediterranean													
ES	2002	35.57	-8.13	0.45	268	21.61	-17.71	0.33	431	21.61	-17.71	0.33	431
	2012	22.85	-12.59	0.39	690	22.72	-12.76	0.39	694	22.91	-12.57	0.39	709
IL	2002	34.18	-13.17	0.42	219	27.72	-14.6	0.4	289	27.76	-14.28	0.4	305
	2012	28.59	-15.24	0.39	298	27.36	-14.88	0.39	299	28.05	-14.64	0.39	304
PT	2002	40.38	-3.42	0.48	159	28.75	-10.88	0.39	185	31.89	-7.95	0.42	289
	2012	30.27	-10.32	0.39	173	29.99	-10.6	0.39	173	30.81	-9.96	0.4	185

Notes: ISSP module "Changing family and gender roles". The absolute hours correspond to raw data subject to imputation procedures described in Appendix B. The gap denotes the average value of the difference between female and male hours in a couple. Relative hours is the average value of the ratio of hours devoted to a particular activity by a female divided by the household total (sum of male and female input). The summary statistics are reported for respondents in the regression sample (married, non-retirement age, with key variables non-missing). Some values for Bulgaria and Great Britain are missing in 2012 because there is no information about spouse's hours worked for pay in the 2012 wave.

Table B.3: Country-level summary statistics for key regression variables by respondent's gender in 2002-2012

Country	Female income share		Female rel. input to market work		Female rel. input to home work		Female rel. input to family care		Agree with traditional division		Agree that both spouses should work	
	F	M	F	M	F	M	F	M	F	M	F	M
2002												
DE-E	0.44	0.36	0.49	0.36	0.7	0.71	.	.	0.1	0.11	0.92	0.86
BG	0.48	0.38	0.53	0.43	0.7	0.67	.	.	0.32	0.5	0.9	0.85
HU	0.41	0.46	0.39	0.43	0.75	0.73	.	.	0.27	0.41	0.81	0.72
LV	0.42	0.38	0.43	0.35	0.65	0.63	.	.	0.42	0.5	0.84	0.71
PL	0.45	0.34	0.59	0.29	0.65	0.62	.	.	0.35	0.41	0.81	0.74
RU	0.15	0.8	0.48	0.42	0.66	0.64	.	.	0.5	0.63	0.85	0.79
SI	0.48	0.44	.	.	0.71	0.73	.	.	0.23	0.24	0.92	0.85
SK	0.39	0.42	.	.	0.69	0.66	.	.	0.42	0.5	0.88	0.82
CZ	0.39	0.39	.	.	0.73	0.69	.	.	0.43	0.55	0.85	0.81
US	0.35	0.34	0.42	0.33	0.7	0.64	.	.	0.17	0.22	0.54	0.54
GB	0.32	0.31	0.36	0.33	0.73	0.64	.	.	0.06	0.12	0.46	0.48
DK	0.43	0.39	0.46	0.44	0.68	0.62	.	.	0.06	0.07	0.75	0.71
FI	0.43	0.4	0.43	0.43	0.69	0.65	.	.	0.03	0.11	0.67	0.63
NO	0.38	0.35	0.36	0.43	0.75	0.7	.	.	0.02	0.08	0.64	0.7
SE	0.45	0.41	0.44	0.42	0.65	0.62	.	.	0.02	0.05	0.83	0.82
FR	0.44	0.35	0.42	0.37	0.8	0.69	.	.	0.08	0.19	0.75	0.67
DE-W	0.29	0.24	0.33	0.2	0.73	0.76	.	.	0.11	0.19	0.67	0.5
AT	0.39	0.3	0.41	0.35	0.77	0.75	.	.	0.2	0.36	0.83	0.72
ES	0.51	0.22	0.42	0.28	0.76	0.74	.	.	0.08	0.21	0.92	0.87
IL	0.44	0.3	0.46	0.34	0.73	0.72	.	.	0.13	0.32	0.85	0.77
PT	0.47	0.3	0.41	0.43	0.82	0.75	.	.	0.22	0.22	0.98	0.92
2012												
DE-E	0.4	0.37	0.43	0.37	0.74	0.66	0.62	0.57	0.03	0.03	0.83	0.89
BG	0.35	0.46	.	.	0.73	0.72	0.67	0.63	0.37	0.31	0.94	0.92
HU	0.37	0.49	0.4	0.42	0.76	0.69	0.63	0.66	0.41	0.45	0.7	0.82
LV	0.41	0.38	0.44	0.39	0.67	0.62	0.63	0.59	0.57	0.51	0.79	0.74
PL	0.28	0.53	0.36	0.45	0.68	0.62	0.66	0.6	0.31	0.45	0.81	0.74
RU	0.38	0.41	0.44	0.39	0.67	0.64	0.68	0.67	0.49	0.53	0.75	0.84
SI	0.44	0.43	0.46	0.38	0.75	0.72	0.61	0.55	0.12	0.19	0.92	0.92
SK	0.42	0.42	0.45	0.45	0.71	0.63	0.66	0.61	0.46	0.58	0.84	0.78
CZ	0.38	0.44	0.4	0.44	0.72	0.67	0.64	0.62	0.46	0.51	0.89	0.91
US	0.36	0.36	0.35	0.35	0.67	0.64	0.56	0.7	0.19	0.2	0.5	0.59
GB	0.26	0.49	.	.	0.67	0.63	0.64	0.58	0.06	0.09	0.56	0.56
DK	0.48	0.4	0.48	0.44	0.64	0.59	0.6	0.55	0.05	0.05	0.8	0.73
FI	0.46	0.39	0.46	0.46	0.66	0.61	0.64	0.55	0.06	0.11	0.77	0.76
NO	0.5	0.37	0.46	0.43	0.69	0.59	0.61	0.54	0.02	0.05	0.81	0.78
SE	0.45	0.4	0.45	0.41	0.62	0.58	0.57	0.55	0.02	0.08	0.85	0.87
FR	0.44	0.36	0.43	0.4	0.74	0.65	0.68	0.56	0.04	0.07	0.8	0.74
DE-W	0.27	0.27	0.3	0.33	0.75	0.71	0.66	0.61	0.16	0.15	0.71	0.61
AT	0.41	0.36	0.42	0.34	0.73	0.68	0.61	0.6	0.26	0.37	0.76	0.67
ES	0.4	0.32	0.4	0.38	0.73	0.69	0.63	0.59	0.06	0.13	0.97	0.89
IL	0.41	0.39	0.37	0.43	0.75	0.68	0.67	0.59	0.15	0.18	0.83	0.86
PT	0.65	0.28	0.42	0.38	0.79	0.71	0.64	0.61	0.13	0.19	0.98	0.96

Notes: ISSP module "Changing family and gender roles". No survey weights are applied. Except for Bulgaria, Slovenia, Slovakia, Czechia, and Great Britain, the summary statistics are reported for the regression sample (married respondents of non-retirement age with non-missing information about all key variables). The specified countries are excluded from regression analysis because the hours worked by the spouse is not available, and, therefore, relative female input to market work is not possible to compute. The information about hours worked for pay is subject to the "upper bound" imputation procedure described in Appendix B.

Table B.4: Married women's time allocation by country in 2002-2012

Country	Year	Market work Abs., hours	Market work Gap, hours	Market work Rel., share	Household work Abs., hours	Household work Gap, hours	Household work Rel., share	Family care Abs., hours	Family care Gap, hours	Family care Rel., share	Task sharing index	Income share	N
Post-socialist													
DE-E	2002	28.89	-10.2	0.41	16.78	10.32	0.71	.	.	.	2.17	0.4	116
	2012	28.05	-12.44	0.4	16.29	9.56	0.7	13.39	5.89	0.59	2.21	0.38	123
HU	2002	26.97	-11.55	0.41	27.37	17.05	0.74	.	.	.	1.93	0.44	273
	2012	26.4	-7.8	0.41	25.03	15.09	0.72	22.14	11.13	0.64	2.1	0.44	102
BG	2002	29.46	-2.92	0.48	22.23	11.67	0.69	.	.	.	1.97	0.43	315
	2012	28.51	.	.	21.8	12.95	0.73	18.62	10.26	0.65	2.05	0.4	146
LV	2002	31.64	-12.74	0.39	19.16	7.39	0.64	.	.	.	2.19	0.4	278
	2012	30.42	-8.12	0.42	23.6	10.04	0.65	16.92	6.96	0.61	2.18	0.39	227
PL	2002	34.49	-9.07	0.44	19.99	7.92	0.64	.	.	.	2.09	0.39	333
	2012	28.11	-12.04	0.4	24.74	9.62	0.65	22.25	11.53	0.63	1.97	0.39	223
RU	2002	34.48	-5.44	0.45	24.3	10.33	0.65	.	.	.	2.19	0.45	324
	2012	30.01	-10.4	0.42	27.08	12.03	0.66	22.22	12.59	0.67	1.94	0.4	308
SI	2002	35.87	.	.	20.14	12.1	0.72	.	.	.	2.13	0.46	107
	2012	29.36	-8.5	0.42	24.21	15.11	0.73	16.94	6.5	0.58	2.2	0.43	236
SK	2002	34.84	.	.	20.96	9.95	0.67	.	.	.	2.08	0.4	253
	2012	30.42	-5.98	0.45	21.37	9.71	0.67	16.69	8.49	0.63	2.05	0.42	255
CZ	2002	31.02	.	.	22.72	13.1	0.72	.	.	.	1.89	0.39	221
	2012	34.12	-9.14	0.42	19.58	10.21	0.69	15.64	7.84	0.63	1.97	0.41	422
Liberal													
US	2002	29.28	-14.33	0.38	13.3	7.08	0.67	.	.	.	2.38	0.34	297
	2012	24.53	-16.13	0.35	16.91	8.73	0.66	34.09	14.67	0.63	2.28	0.36	322
GB	2002	24.06	-17.68	0.35	12.97	7.12	0.69	.	.	.	2.26	0.31	341
	2012	20.48	.	.	12.14	5.56	0.65	27.9	12.32	0.6	2.32	0.39	84
Nordic													
DK	2002	33.52	-7.12	0.45	12.72	5.47	0.65	.	.	.	2.33	0.41	424
	2012	32.19	-5.09	0.46	11.81	4.25	0.61	15.77	4.34	0.58	2.44	0.44	429
FI	2002	28.9	-7.6	0.43	11.51	5.85	0.67	.	.	.	2.39	0.41	475
	2012	32.92	-5.93	0.46	11.61	4.82	0.63	23.07	9.12	0.59	2.44	0.42	290
NO	2002	29.12	-12.44	0.39	11.33	6.91	0.72	.	.	.	2.3	0.36	457
	2012	34.01	-6.9	0.44	10.92	4.28	0.64	17.72	5.23	0.57	2.44	0.43	420
SE	2002	31.94	-8.07	0.43	13.19	5.75	0.63	.	.	.	2.45	0.43	407
	2012	33.4	-7.6	0.43	13.14	4.08	0.6	17.91	4.26	0.56	2.55	0.43	239
Conservative													
FR	2002	28.31	-11.36	0.4	11.91	8.31	0.76	.	.	.	2.08	0.41	552
	2012	29.13	-9.74	0.42	10.69	6.12	0.71	20.69	8.9	0.64	2.12	0.41	541
DE-W	2002	16.42	-25.24	0.25	21.68	14.95	0.75	.	.	.	2.09	0.26	219
	2012	20.79	-19.77	0.32	19	12.36	0.73	20.7	10.56	0.63	2.11	0.27	326
AT	2002	24.64	-13.67	0.38	21.57	14.81	0.76	.	.	.	1.97	0.34	390
	2012	23.72	-12.96	0.38	19.94	12.2	0.71	18.39	10	0.61	2.35	0.39	286
Mediterranean													
ES	2002	21.61	-17.71	0.33	26.87	18.71	0.75	.	.	.	2.13	0.32	431
	2012	22.91	-12.57	0.39	26.21	16.28	0.71	25.59	10.66	0.61	2.27	0.36	709
IL	2002	27.76	-14.28	0.4	16.17	10.31	0.72	.	.	.	2.22	0.37	305
	2012	28.05	-14.64	0.39	20.25	12.8	0.72	23.18	11.13	0.64	2.26	0.41	304
PT	2002	31.89	-7.95	0.42	21.84	16.41	0.79	.	.	.	2.02	0.39	289
	2012	30.81	-9.96	0.4	21.82	14.13	0.75	13.16	6.57	0.62	2.22	0.46	185

Notes: ISSP module "Changing family and gender roles". No survey weights are applied. The sample for which summary statistics are reported is provided in the column title.

Table B.5: OLS regression results for determinants of share of time devoted to household work by married women in 2002-2012

	Panel A: All respondents				Panel B: Female respondents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
=1 if 2012	-0.03***	-0.03***	-0.03***	-0.03***	-0.02***	-0.02***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
=1 if post-socialist (PS)	-0.06*	-0.09***	-0.06*	-0.03	-0.07***	-0.10***	-0.08***	-0.06**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
PS × 2012	0.04***	0.04***	0.04***	0.04***	0.05***	0.05***	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\widetilde{h^{fem}_{mkt}}$	-0.15***	-0.17***	-0.15***	-0.15***	-0.15***	-0.17***	-0.15***	-0.15***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\widetilde{h^{fem}_{mkt}} \times PS$		0.08***				0.06***		
		(0.02)				(0.02)		
$\widetilde{y^{fem}_{mkt}}$	-0.06***	-0.06***	-0.06***	-0.06***	-0.04***	-0.04***	-0.05***	-0.04***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
$\widetilde{y^{fem}_{mkt}} \times PS$			0.02				0.03	
			(0.02)				(0.03)	
=1 if traditional division	0.02***	0.02***	0.02***	0.03***	0.02***	0.02***	0.02***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
trad.division × PS				-0.04***				-0.04***
				(0.01)				
=1 if both should work	-0.02***	-0.02***	-0.02***	-0.01**	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
both work × PS				-0.01				-0.01
				(0.01)				(0.01)
=1 if college educated	-0.03***	-0.03***	-0.03***	-0.03***	-0.04***	-0.04***	-0.04***	-0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
# children								
≤ 6 y.o	0	0	0	0	0	0	0	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
> 6 y.o	0.01**	0.01**	0.01**	0.01**	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
cohort effect								
< 10 y.o in 1989	-0.02**	-0.02**	-0.02**	-0.02**	-0.04***	-0.04***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
10-30 y.o in 1989	-0.01*	-0.01	-0.01*	-0.01*	-0.02***	-0.02***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
religion								
=1 if no religion	-0.05***	-0.05***	-0.05***	-0.05***	-0.03***	-0.03***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
=1 if Protestant	-0.05***	-0.05***	-0.05***	-0.05***	-0.06***	-0.06***	-0.06***	-0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
=1 if Christian Orthodox	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
settlement type								
=1 if suburban or town	0.01	0.01	0.01	0.01	0	0	0	0
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
=1 if rural	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
log(family income)	0	0	0	0	0	0	0	0
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	10,604	10,604	10,604	10,604	5,454	5,454	5,454	5,454
R-squared	0.11	0.12	0.11	0.11	0.12	0.12	0.12	0.12
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered on country level estimated using bootstrap are in parentheses: *** p<0.01, ** p<0.05, * p<0.1. No weights are applied. The descriptive statistics are provided in Table 2.2 and 2.3. Notation: $\widetilde{h^{fem}_{mkt}}$ – share of time devoted by a female household member to work for pay in household total. $\widetilde{y^{fem}_{mkt}}$ – share of female income in household income. "=1 if traditional division" – respondent agrees with a claim that "Men's job is to earn money and women's job is to look after home and family". " =1 if both should work" means that respondent agrees with a claim that "Both husband and wife should contribute to household income". "< 10 y.o in 1989" and "10-30 y.o in 1989" capture the cohort effect based on how much time people have spent under socialism – less than 10 year (i.e., born after 1979), between 10 and 30 years (i.e., born between 1959 and 1979) or more than 30 years (omitted option). The omitted category is Roman Catholic for religion and urban for settlement urban-rural status.

Table B.6: OLS regression results for determinants of share of time devoted to family care by married women in 2012

	Panel A: All respondents				Panel B: Female respondents			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
=1 if post-socialist (PS)	0.02	0.01	0.02	0.05***	0.01	-0.01	0.02	0.05
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)
$\widetilde{h_{mkt}^{fem}}$	-0.13***	-0.13***	-0.13***	-0.13***	-0.14***	-0.15***	-0.14***	-0.14***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
$\widetilde{h_{mkt}^{fem}} \times PS$		0.02				0.05		
		(0.04)				(0.06)		
$\widetilde{y_{mkt}^{fem}}$	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\widetilde{y_{mkt}^{fem}} \times PS$			-0.02				-0.02	
			(0.04)				(0.05)	
=1 if traditional division	0.02*	0.02*	0.02*	0.02*	0.02	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
trad.division × PS				-0.01				0.00
				(0.02)				(0.04)
=1 if both should work	-0.02**	-0.02**	-0.02**	-0.01	-0.01	-0.01	-0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
both should work × PS				-0.04***				-0.04*
				(0.02)				(0.02)
=1 if college educated	-0.01**	-0.01**	-0.01**	-0.01**	-0.02*	-0.02*	-0.02*	-0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
# children								
$\leq$ 6 y.o	0.02***	0.02***	0.02***	0.02***	0.03***	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
> 6 y.o	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
cohort effect								
< 10 y.o in 1989	-0.02*	-0.02*	-0.02*	-0.02*	-0.04**	-0.04**	-0.04**	-0.04**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
10-30 y.o in 1989	-0.01	-0.01	-0.01	-0.01	-0.03	-0.03	-0.03	-0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
religion								
=1 if no religion	-0.02***	-0.02***	-0.02***	-0.02***	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
=1 if Protestant	-0.02*	-0.02*	-0.02*	-0.02	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
=1 if Christian Orthodox	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
settlement type								
=1 if suburban or town	0.00				0.00	0.00	0.00	
	(0.01)	(0.01)	(0.01)		(0.01)	(0.01)		(0.01)
=1 if rural	0.02*	0.02*	0.02*	0.02*	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
log(family income)	0	0.00	0.00	0.00	0.00	0.00	-0.01	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	4,829	4,829	4,829	4,829	2,542	2,542	2,542	2,542
R-squared	0.06	0.06	0.06	0.06	0.08	0.08	0.08	0.08
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered on country level estimated using bootstrap are in parentheses: *** p<0.01, ** p<0.05, * p<0.1. No weights are applied. The descriptive statistics are provided in Table 2.2 and 2.3. Notation: $\widetilde{h_{mkt}^{fem}}$ – share of time devoted by a female household member to work for pay in household total. $\widetilde{y_{mkt}^{fem}}$ – share of female income in household income. "=1 if traditional division" – respondent agrees with a claim that "Men's job is to earn money and women's job is to look after home and family". " =1 if both should work" means that respondent agrees with a claim that "Both husband and wife should contribute to household income". "< 10 y.o in 1989" and "10-30 y.o in 1989" capture the cohort effect based on how much time people have spent under socialism – less than 10 year (i.e., born after 1979), between 10 and 30 years (i.e., born between 1959 and 1979) or more than 30 years (omitted option). The omitted category is Roman Catholic for religion and urban for settlement urban-rural status.

Table B.7: Kitagawa-Blinder-Oaxaca decomposition of share of time devoted by married women to different activities by policy cluster in 2002-2012

	Panel A: All respondents				Panel B: Female respondents			
	Liberal	Nordic	Conser-vative	Mediter-ranean	Liberal	Nordic	Conser-vative	Mediter-ranean
Part 1: Household work								
2002								
Mean, advanced	0.674	0.671	0.75	0.76	0.709	0.697	0.783	0.774
Mean, post-socialist	0.67	0.67	0.67	0.67	0.678	0.678	0.678	0.678
Difference in means	**0.004**	**0.001**	**0.08**	**0.091**	**0.031**	**0.02**	**0.106**	**0.096**
	[.016]	[.022]	[.02]	[.024]	[.016]	[.022]	[.025]	[.026]
Endowment effect	**0.079**	**0.072**	**0.007**	**0.001**	**0.048**	**0.046**	**0.01**	**0.025**
	[.016]	[.025]	[.017]	[.021]	[.018]	[.024]	[.016]	[.028]
Determinants	*0.017*	*0.002*	*0.011*	*0.005*	*0.01*	*0.011*	*0.013*	*0.011*
	[.004]	*[.004]*	*[.007]*	*[.004]*	*[.01]*	*[.009]*	*[.008]*	*[.009]*
Coefficient effect	**0.01**	**-0.029**	**0.105**	**0.078**	**0.063**	**0.035**	**0.125**	**0.123**
	[.031]	[.026]	[.052]	[.049]	[.039]	[.027]	[.071]	[.055]
Determinants	*-0.031*	*-0.035*	*0.002*	*-0.013*	*-0.054*	*-0.06*	*-0.02*	*-0.074*
	[.018]	*[.024]*	*[.019]*	*[.025]*	*[.032]*	*[.044]*	*[.032]*	*[.038]*
Interaction effect	**-0.085**	**-0.042**	**-0.031**	**0.012**	**-0.08**	**-0.061**	**-0.03**	**-0.052**
	[.031]	[.028]	[.053]	[.053]	[.04]	[.021]	[.068]	[.068]
2012								
Mean, advanced	0.684	0.607	0.706	0.725	0.747	0.661	0.728	0.74
Mean, post-socialist	0.66	0.66	0.66	0.66	0.72	0.72	0.72	0.72
Difference in means	**0.024**	**-0.053**	**0.046**	**0.065**	**0.027**	**-0.058**	**0.008**	**0.02**
	[.013]	[.016]	[.018]	[.017]	[.027]	[.03]	[.029]	[.028]
Endowment effect	**0.052**	**0.035**	**-0.002**	**-0.013**	**0.034**	**0.018**	**-0.008**	**-0.036**
	[.017]	[.023]	[.011]	[.011]	[.031]	[.042]	[.011]	[.021]
Determinants	*0.013*	*-0.005*	*0.007*	*0*	*0*	*-0.019*	*-0.008*	*-0.015*
	[.003]	*[.003]*	*[.006]*	*[.004]*	*[.005]*	*[.004]*	*[.007]*	*[.007]*
Coefficient effect	**0.047**	**-0.056**	**0.029**	**0.061**	**0.002**	**-0.063**	**-0.022**	**0.098**
	[.016]	[.042]	[.02]	[.033]	[.034]	[.045]	[.032]	[.07]
Determinants	*0.019*	*0.012*	*0.014*	*-0.027*	*0.004*	*-0.034*	*-0.037*	*-0.049*
	[.012]	*[.019]*	*[.02]*	*[.032]*	*[.014]*	*[.025]*	*[.016]*	*[.042]*
Interaction effect	**-0.075**	**-0.032**	**0.019**	**0.018**	**-0.009**	**-0.014**	**0.038**	**-0.043**
	[.019]	[.044]	[.014]	[.023]	[.038]	[.051]	[.019]	[.063]
Part 2: Family care								
2012								
Mean, advanced	0.687	0.577	0.619	0.615	0.509	0.627	0.666	0.666
Mean, post-socialist	0.634	0.634	0.634	0.634	0.653	0.653	0.653	0.653
Difference in means	**0.053**	**-0.057**	**-0.015**	**-0.019**	**-0.144**	**-0.026**	**0.013**	**0.013**
	[.014]	[.016]	[.016]	[.017]	[.018]	[.037]	[.03]	[.022]
Endowment effect	**0.027**	**0.023**	**0.001**	**0.013**	**0.014**	**-0.003**	**0.006**	**0.041**
	[.02]	[.03]	[.013]	[.036]	[.031]	[.048]	[.013]	[.064]
Determinants	*0.016*	*-0.009*	*0.004*	*-0.008*	*0.01*	*-0.018*	*-0.006*	*-0.015*
	[.005]	*[.006]*	*[.008]*	*[.006]*	*[.012]*	*[.012]*	*[.011]*	*[.011]*
Coefficient effect	**-0.015**	**-0.008**	**-0.04**	**-0.049**	**-0.114**	**0.019**	**0.006**	**-0.011**
	[.018]	[.023]	[.025]	[.03]	[.053]	[.037]	[.03]	[.028]
Determinants	*-0.032*	*0.069*	*0.039*	*0.024*	*-0.11*	*0.071*	*0.007*	*-0.015*
	[.014]	*[.032]*	*[.017]*	*[.015]*	*[.033]*	*[.04]*	*[.042]*	*[.043]*
Interaction effect	**0.041**	**-0.072**	**0.025**	**0.017**	**-0.044**	**-0.042**	**0.001**	**-0.018**
	[.023]	[.036]	[.025]	[.041]	[.059]	[.055]	[.031]	[.056]

Notes: The regression equation used as a basis for decomposition includes all the variables used in column 1 of Appendix Table B.5 and B.6. Determinants stand for relative female input to market work, household income share contributed by women, and indicator variables for gender attitudes. The decomposition was performed using the oaxaca module in Stata (Jann et al., 2008).

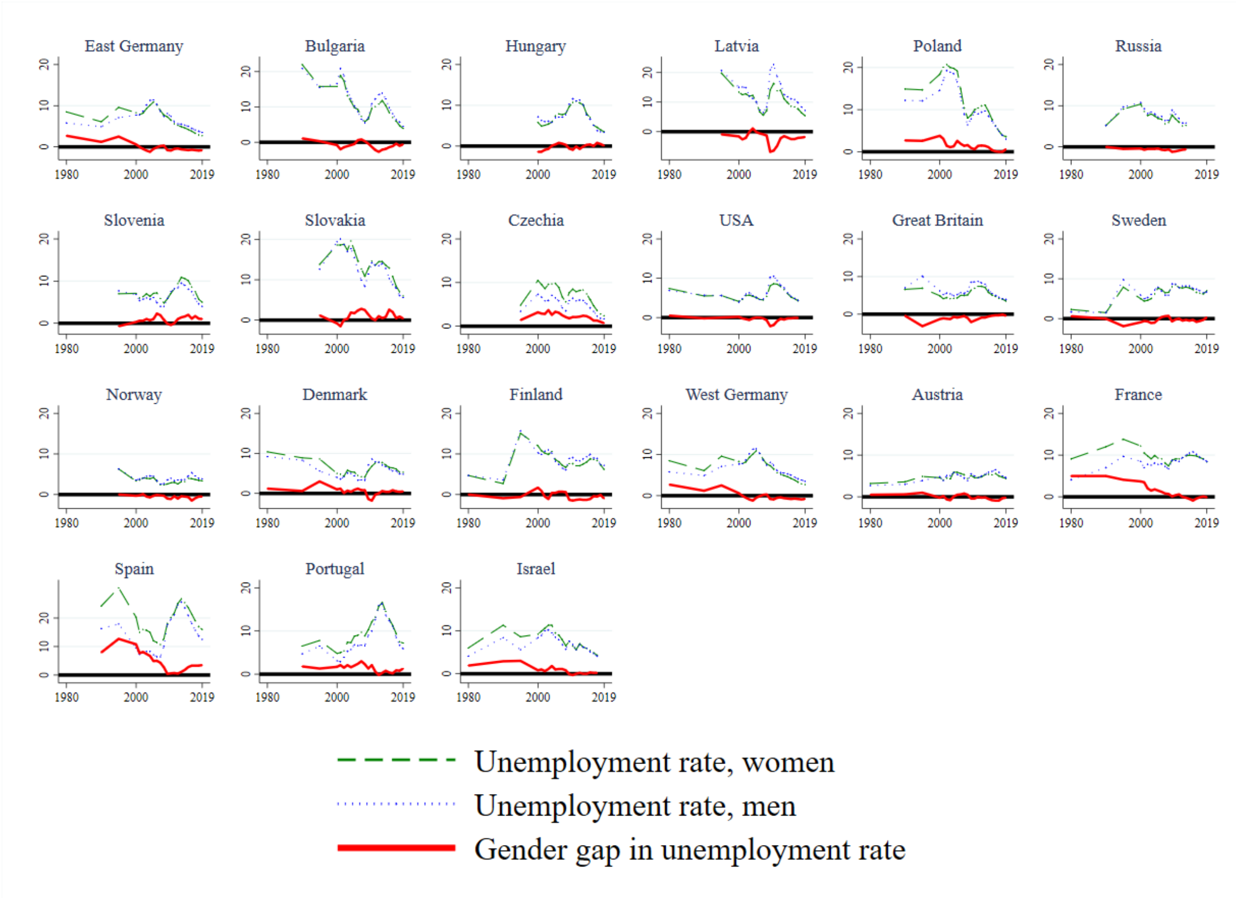


Figure B.1: Aggregate unemployment rate by country and gender

Notes: The graphs plot aggregate unemployment rates according to the UNECE data. The gender gap is defined as the difference between male and female unemployment rate (i.e., positive value of the gap means that female unemployment rate exceeds male unemployment rate).

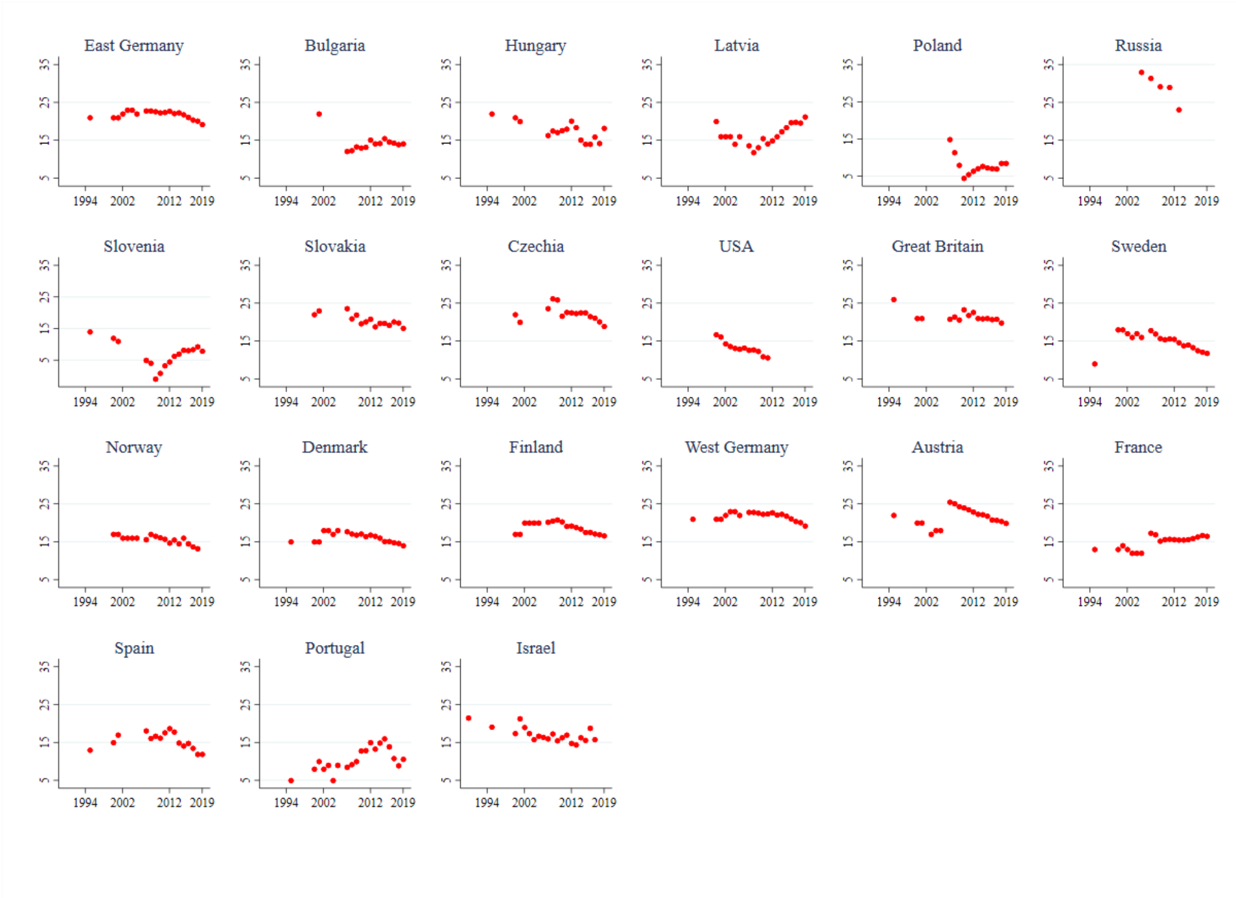


Figure B.2: Aggregate gender pay gap in hourly earnings, %

Notes: The graphs plot aggregate gender pay gap in hourly earnings according to the UNECE data. Gender pay gap is the difference between men's and women's average earnings from employment, shown as a percentage of men's average earnings. Gender pay gap in hourly earnings aims to capture the difference between men's and women's overall position in the labor market. It does not account for the number of hours worked, the type of activity, or the type of occupation.

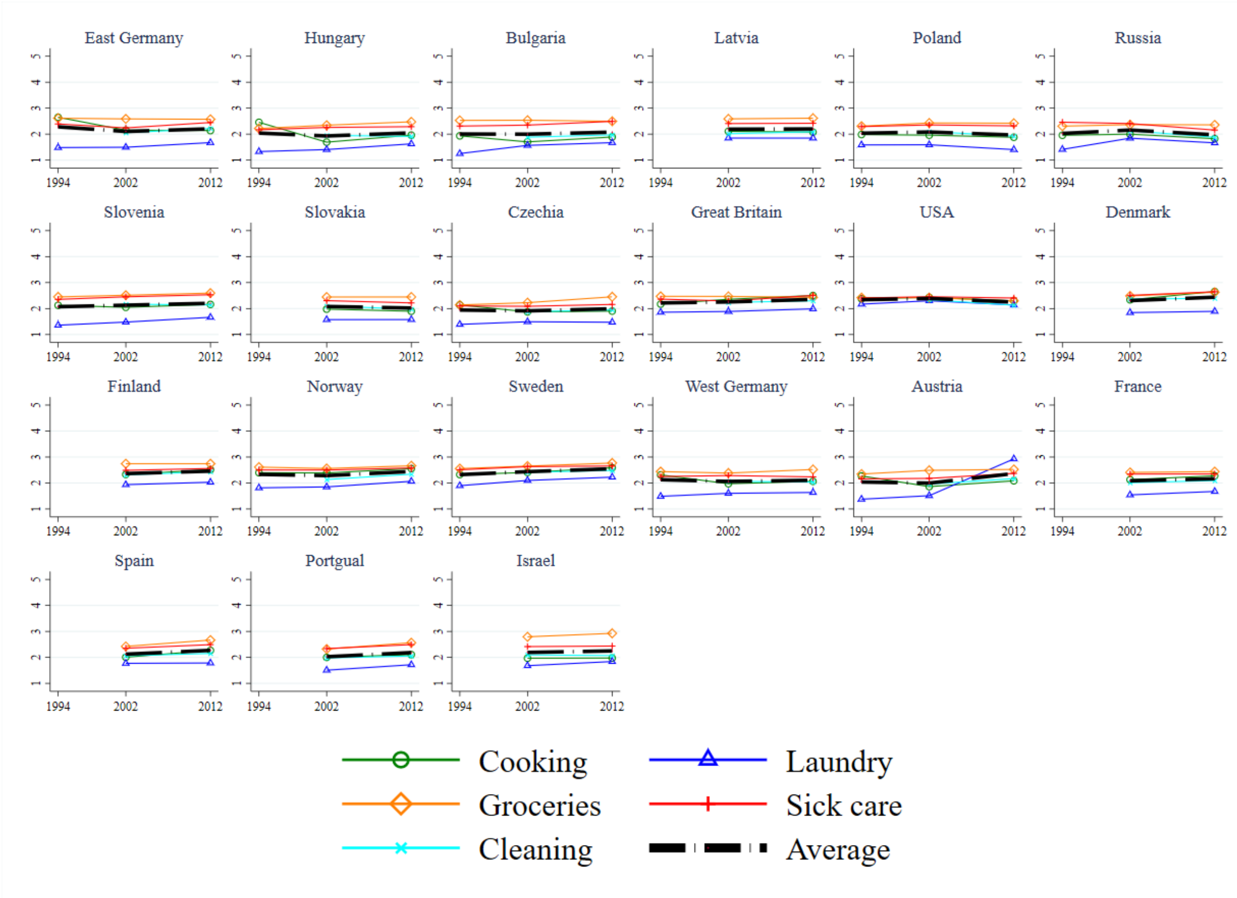


Figure B.3: Allocation of household activities by country (extensive margin)

Notes: The graphs plot the mean of categorical variables where 1 = always woman, 2 = usually woman, 3 = split equally or third person, 4 = usually man, 5 = always man. Survey weights are applied if available.

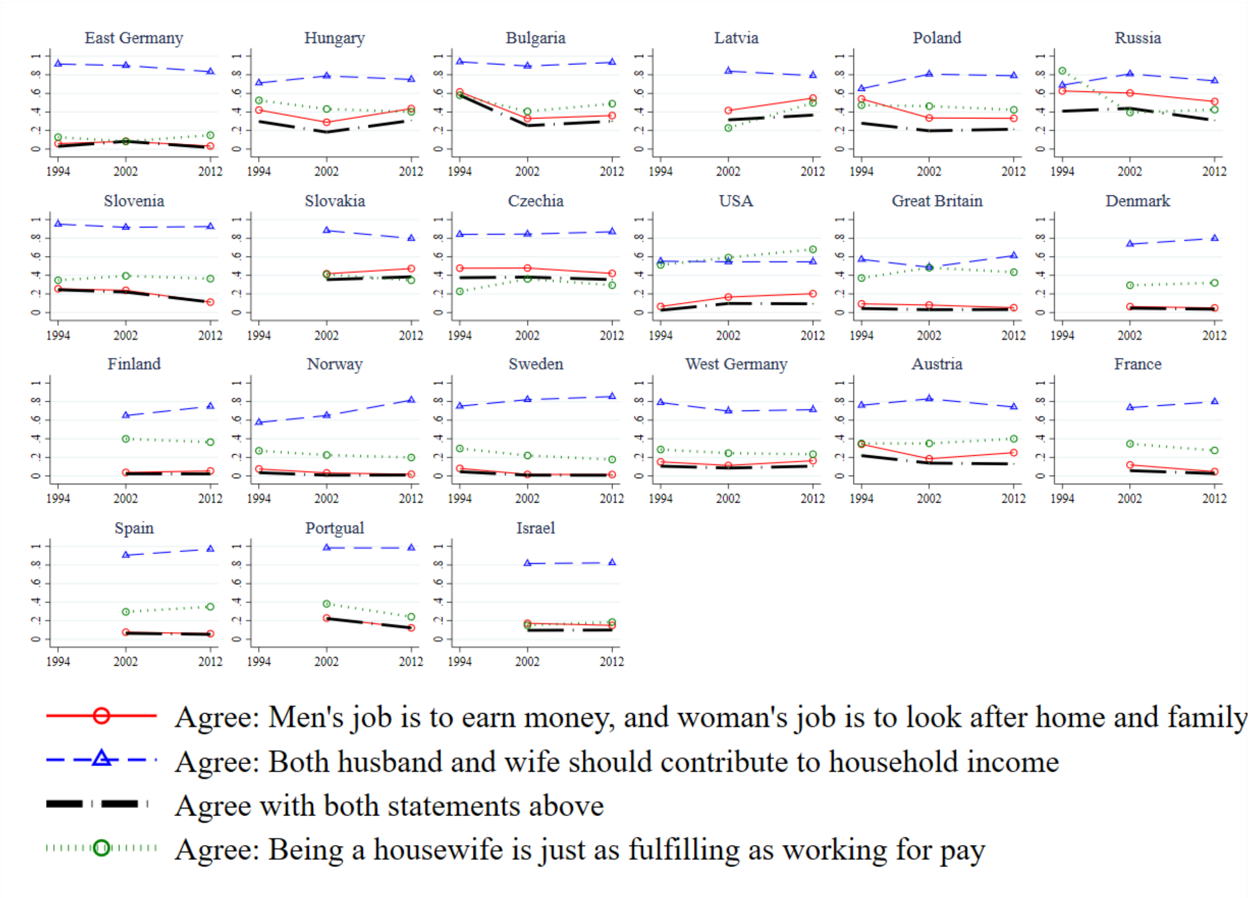


Figure B.4: Gender attitudes of married women by country

Notes: The graphs plot the share of married women of non-retirement age agreeing with a claim in the legend. Survey weights are applied if available.

Appendix C

Appendix for Chapter 3

C.1 Descriptive Statistics

C.1.1 Attrition

When we launched our first wave of the survey, 10,758 MTurk workers attempted to participate in our survey. Among then 5,487 MTurkers completed the first wave of the survey. We examine if the attrition is systematically correlated with treatment arms. Table C.1 shows that the attrition rates are not different across treatment arms.

Table C.1: Attrition rates by treatment arms ($N = 10,758$)

CPI	Wage	Unemp	AQI	Vax
0.50	0.50	0.48	0.49	0.48

To further examine if the attrition is systematically different across treatment arms, we regress the indicator variable denoting the attrition on treatment arm dummies. Table C.2 further illustrates that attrition is not systematically related to the treatment arms.

Table C.2: Regression of attrition rates on treatment arms

treat_cpi	treat_unemp	treat_vax	treat_wave	Constant
0.007	-0.011	-0.011	0.013	0.489***
(0.015)	(0.016)	(0.015)	(0.014)	(0.010)

Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.3 below summarizes the attrition from participating in the follow-up surveys. It shows that attrition was the highest in the control group that received the information about the air quality index in Seattle. The attrition rates between the two treatment groups are similar. This likely happened because workers might have found the information about the air quality in Seattle less interesting than the one about CPI or hourly earnings inflation rates. Another reason might be that air quality transcription task asked workers to record four numbers rather than three as is the case for the treatment groups (CPI and Wage groups). We also find that, overall, older workers and those without children are more likely to participate in the follow-up waves. Other than this, there are no systematic differences for other demographic characteristics.

Table C.3: Attrition rates from participating in the follow-up waves

Wave 1 $\rightarrow$ Wave2			Wave 1 $\rightarrow$ Wave 3			All three waves		
CPI	Wage	AQI	CPI	Wage	AQI	CPI	Wage	AQI
0.53	0.51	0.62	0.46	0.44	0.59	0.65	0.63	0.72

C.1.2 Descriptive statistics (follow-up surveys)

Table C.4 below provides descriptive statistics about respondents who participated in the second and third waves. Table C.4 shows that they are similar to those from the first wave of the survey in Section 3.2.3.

Table C.4: Descriptive Statistics (Wave 2&3)

Wave 2 (June 2022)	Mean	Percentiles			Std. Dev.
		p25	p50	p75	
age	40.51	31.00	39.00	49.00	12.26
female	0.48	0.00	0.00	1.00	0.50
white	0.80	0.00	1.00	1.00	0.40
with college degree	0.74	0.00	1.00	1.00	0.44
employed	0.82	0.00	1.00	1.00	0.39
full-time employed	0.69	0.00	1.00	1.00	0.46
number of children	0.85	0.00	1.00	2.00	1.02
monthly spending on food	\$588.65	\$185.00	\$350.00	\$600.00	2259.52
monthly spending on gas	\$402.35	\$50.00	\$100.00	\$200.00	7855.65
$\mathbb{E}_t^{\texttt{prior}}[\pi_{t+12}]$	5.61	1.00	5.00	10.00	8.18
$\mathbb{E}_t^{\texttt{prior}}[\pi_{t+12}^{w}]$	5.70	1.00	5.00	8.00	9.87
$\mathbb{E}_t^{\texttt{prior}}[u_{t+12}]$	7.06	4.31	6.30	9.00	3.55
$\mathbb{E}_t^{\texttt{prior}}[\texttt{duration}_{t+1}]$	3.87	3.00	5.00	5.00	1.49
$\mathbb{E}_t^{\texttt{prior}}[\texttt{reservation wages}_{t+1}]$	0.94	0.50	0.92	1.17	0.54
Observations	1,460				

Wave 3 (July 2022)	Mean	Percentiles			Std. Dev.
		p25	p50	p75	
age	40.78	31.00	39.00	49.00	12.22
female	0.49	0.00	0.00	1.00	0.50
white	0.81	0.00	1.00	1.00	0.39
with college degree	0.74	0.00	1.00	1.00	0.44
employed	0.82	0.00	1.00	1.00	0.38
full-time employed	0.67	0.00	1.00	1.00	0.47
number of children	0.89	0.00	1.00	2.00	1.09
monthly spending on food	\$519.85	\$150.00	\$350.00	\$560.00	1166.04
monthly spending on gas	\$205.99	\$50.00	\$120.00	\$225.00	361.25
$\mathbb{E}_t^{\texttt{prior}}[\pi_{t+12}]$	5.30	1.00	4.00	9.00	9.42
$\mathbb{E}_t^{\texttt{prior}}[\pi_{t+12}^{w}]$	5.30	1.00	3.00	6.00	9.42
$\mathbb{E}_t^{\texttt{prior}}[u_{t+12}]$	6.97	4.25	6.20	8.90	3.47
$\mathbb{E}_t^{\texttt{prior}}[\texttt{duration}_{t+1}]$	3.98	3.00	5.00	5.00	1.44
$\mathbb{E}_t^{\texttt{prior}}[\texttt{reservation wages}_{t+1}]$	0.99	0.50	1.00	1.25	0.54
Observations	1,470				

C.2 Effects of Information Treatment on Subjective Expectations

This section supplements Section 3.3. First, we present binned scatter plots of respondents' posterior expectations after the information provision against their priors by each treatment (CPI inflation, hourly earnings inflation, unemployment, and all three pooled together). Second, we provide regression results from alternative specifications to study information treatment effects.

C.2.1 Graphical Illustration of Information Treatment Effects

This section presents binned scatter plots of respondents' posterior expectations against their priors by each treatment (CPI inflation, hourly earnings inflation, unemployment, and all three pooled together). Consistent with discussion in Section 3.3, Figure C.1 shows that respondents in the treatment group put smaller weights on their prior when they received the relevant signals, whether it is information about price inflation or other macroeconomic variables. Treatment groups exhibit much flatter slopes in all cases. Respondents adjust their weights towards the signal the most when they have received the information about the CPI inflation.

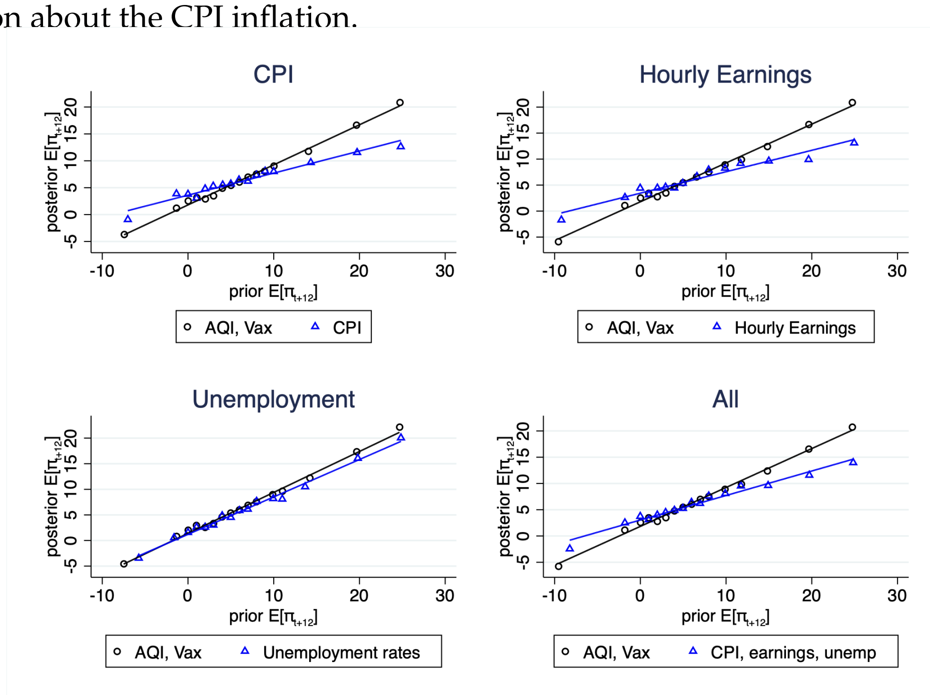
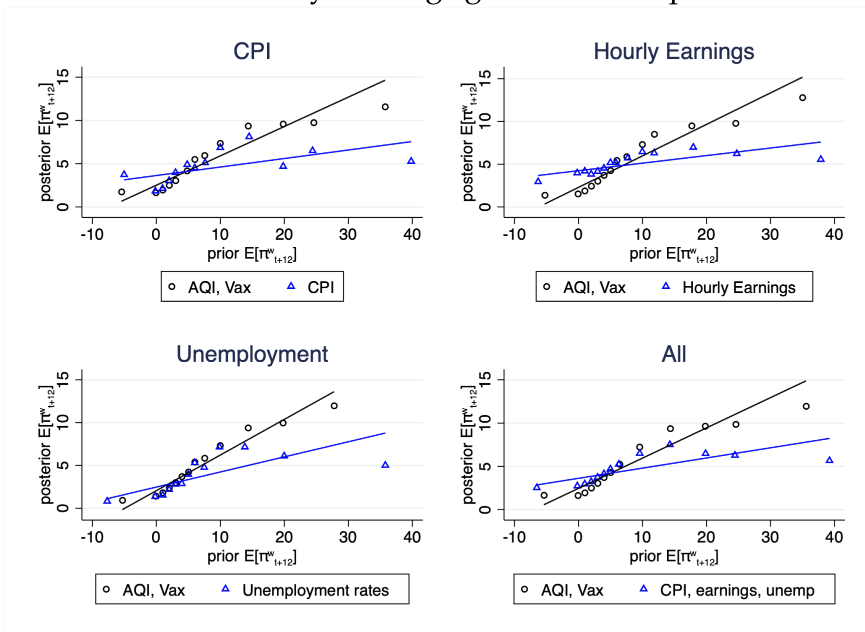


Figure C.1: Effects of information treatment on price inflation expectations

Notes: This figure draws binned scatter plots of highly numerate respondents' posterior expected price inflation rates over the next 12 months (on y-axis) against their priors (on $x-$axis) from the first wave of the survey. Huber-robust weights are applied. Blue triangles are for those who have received the relevant information treatment and black circles are for those who have received irrelevant information about the air quality index (AQI) in Seattle or Covid-19 vaccination rates (Vax). Panels 1-4 refer respectively to CPI inflation treatment, hourly earnings treatment, unemployment rate, and all treatments pooled together.

Figure C.2 paints the same picture. The slopes are much flatter for those in the treatment groups, suggesting that respondents in the treatment group update their expectations about either hourly earnings inflation or unemployment rates after receiving the relevant signals. While hourly earnings inflation expectations are more responsive to the signals about price and hourly earnings inflation, the unemployment rate responds mostly to the signal about unemployment rates. The above figures illustrate the effect of information provision on subjective expectations (price and wage inflation rates and unemployment rates).

Revision of hourly earnings growth rate expectations



Revision of unemployment rate expectations

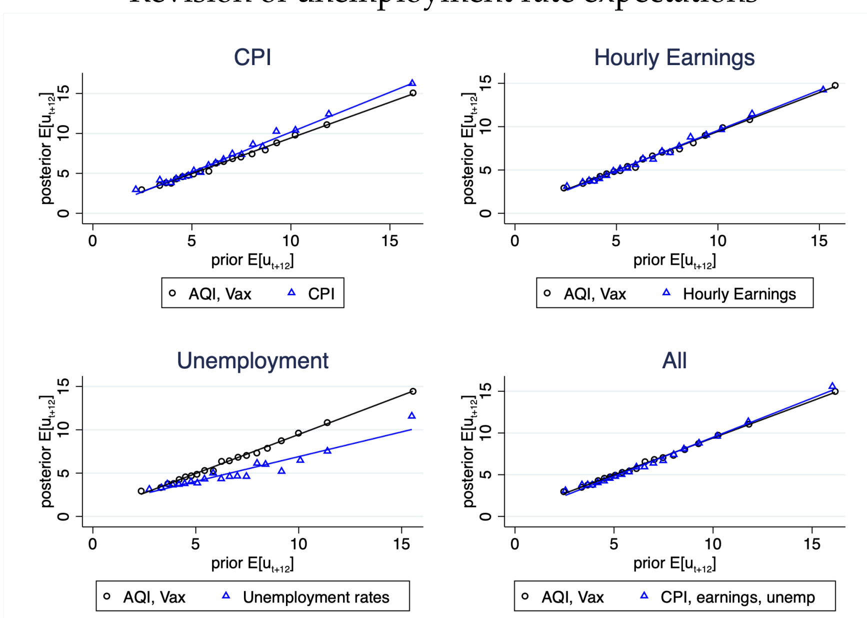


Figure C.2: Effects of information treatment on hourly earnings and unemployment rates expectations

Notes: This figure draws binned scatter plots of highly numerate respondents' posterior expected wage inflation rates (upper panel) and unemployment rates (lower panel) over the next 12 months (on y-axis) against their priors (on $x-$axis) from the first wave of the survey. Huber-robust weights are applied. Blue triangles are for those who have received the relevant information treatment and black circles are for those who have received irrelevant information about the air quality index (AQI) in Seattle or Covid-19 vaccination rates (Vax). Panels 1-4 refer respectively to CPI inflation treatment, hourly earnings treatment, unemployment rate, and all treatments pooled together.

C.2.2 Information Treatment Effects on Broad Regime Changes in Expectations

This section summarizes information treatment effects on broad regime changes in expectations to supplement discussion in Section 3.5.1. We extend the specification estimated there by introducing interaction terms of regime change indicators with prior expectations:

$$\begin{aligned}\texttt{Regime Change}_i^Z =& \beta_0 + \beta_1 \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}] + \sum_{k\in\{\pi,\pi^w,u\}} \beta_{2,k}\texttt{treat}_i^k \\ &+ \sum_{k\in\{\pi,\pi^w,u\}} \beta_{3,k}\left(\texttt{treat}_i^k \times \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]\right) + \varepsilon_i, \ \ Z\in\{\pi,\pi^w,u\},\end{aligned} \tag{C.1}$$

where $\texttt{Regime Change}_i^Z$ denotes if a respondent i revises her *qualitative* assessment about variable Z *upwards*. For instance, if a respondent i thinks that the overall price level will stay the same over the next 12 months, before the treatment, and changes this assessment so that she now thinks the overall price level will increase, after the treatment, then $\texttt{Regime Change}_i^\pi$ takes on the value of one. Similarly, if another respondent thinks that the overall price level will decrease over a year, before the treatment, but changes this assessment to "stay the same," or "increase," after the treatment, then $\texttt{Regime Change}_i^\pi$ equals to one. It will take on the value of zero otherwise. We define $\texttt{Regime Change}_i^{\pi^w}$ similarly. Meanwhile, because unemployment rates are defined differently, we define $\texttt{Regime Change}_i^u$ equals to one as long as respondents raise their unemployment expectations after the treatment and zero otherwise.

Table C.5 shows the results. They are in line with the results in Table 3.8 and broadly consistent with the results for actual revisions in Table 3.3. First, columns 1-3 show the results for broad regime changes in forecast revisions on price inflation expectations. They show that when respondents are provided with either the current CPI inflation rate or the current hourly earnings inflation rates, they are more likely to revise their price inflation expectations upwards, on average. As expected, they are less likely to do so, if their prior expectations are already high. Columns 4-6 show the results for broad regime changes in forecast revisions on wage inflation expectations. Again, they show broadly consistent results with Table 3.3. When they are provided with either the current CPI inflation rates or hourly earnings inflation rates, they are more likely to revise wage inflation expectations upwards. As is the case for the price inflation expectations, they are less likely to do so if their prior wage inflation expectations are high from the beginning. Lastly, columns 7-9 show the results from the unemployment rate expectations. They show that those in the treatment group are *less* likely to revise their unemployment expectations upwards when provided with the current unemployment rates. Consistent with the results in Table 3.3, they are mostly responsive to the current unemployment rate information. Moreover, the higher their prior expected unemployment rate is, the smaller becomes the likelihood of revising their expected unemployment rate upward. Interestingly, but consistent with the results in Table 3.3, the higher their prior expected unemployment rate is, the higher becomes the likelihood of moving to higher unemployment rate regimes when provided with the current CPI inflation rates. This again reflects stagflationary view of the U.S. households.

Table C.5: Information treatment effects on broad regime changes in forecast revisions

Dependent variable:	Price inflation ($Z = \pi$)			Wage inflation ($Z = \pi^w$)			Unemployment rate ($Z = u$)		
`Regime Change`$_i^Z$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
`treat_cpi`	0.10***	0.06***	0.08***	0.05***	0.03*	0.04**	-0.02	-0.03	-0.04
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.06)
`treat_wage`	0.09***	0.07***	0.08***	0.17***	0.15***	0.19***	-0.02	-0.03	-0.05
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.06)
`treat_unemp`	0.00	0.02	0.04**	-0.04*	-0.02	0.00	-0.22***	-0.21***	-0.07
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)	(0.07)
`treat_cpi` $\times$ $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.01***	-0.01***	-0.01***	0.00	-0.00	-0.00	0.01**	0.01**	0.01*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
`treat_wage` $\times$ $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
`treat_unemp` $\times$ $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00	-0.02**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
$\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	3903	3840	2810	3841	3766	2768	3694	3623	2637
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output from equation (3.2) for respondents in all control and treatment groups where the outcome variable is an indicator that respondent revised expectations of the variable in column header upward. For each outcome variable, the first column reports results without controls, the second column adds control variables, and third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.3 Learning Effects

This section explores learning effects of information provision. First, we study the long-run effects of information provision. In specific, we examine if the information treatment effects persist in the subsequent follow-up surveys. Second, we study the learning through survey effects by comparing the treatment effects across the three waves.

C.3.1 Bayesian Learning Effects

In this section, we examine if the information treatment effects are persistent over the next few months. To that end, we run the following regression:

$$\begin{aligned}\Delta\mathbb{E}_{it+j}^{\text{prior}j-\text{prior}1}[Z_{t+12}] =& \beta_0 + \beta_1\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}] \ + \sum_{k\in\{\pi,\pi^w,u\}}\beta_{2,k}\texttt{treat}_i^k \\ &+ \sum_{k\in\{\pi,\pi^w,u\}}\beta_{3,k}\left(\texttt{treat}_i^k \times \mathbb{E}_{it}^{\text{prior}}[Z_{t+12}]\right) + \mathbf{X}_i'\gamma + \varepsilon_i,\ j=\{1,2\}\end{aligned} \tag{C.2}$$

for $Z = \{\pi, \pi^w, u\}$. This is similar to the specification in the main text, equation (3.2), but the dependent variable is now the revisions in *prior* expectations from the first wave to the subsequent follow-up waves.

Table C.6 shows the results. From $\hat{\beta}_{3,k}$'s, it is clear that the information treatment effects persist over, at least, two more months. When respondents update their expectations, they still place some weight on the relevant information they received one or two months ago. The implied weights on the new information are, however, smaller than those from Table 3.3. This is consistent with standard Bayesian learning. As time passes, the information gets more dated and so respondents put less weight on the information that they received a month or two months ago.

Table C.6: Effects of information treatments on revision of price inflation, wage inflation, and unemployment expectations (Wave 2-3)

	Price inflation ($Z = \pi$)			Wage inflation ($Z = \pi^w$)			Unemployment rate ($Z = u$)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Part 1: Dependent variable: $\Delta\mathbb{E}_{it}^{\texttt{prior2-prior1}}[Z_{t+12}]$									
`treat_cpi`	0.12	-0.05	0.18	0.49	0.31	0.35	-0.24	-0.12	-0.02
	(0.56)	(0.56)	(0.64)	(0.35)	(0.38)	(0.44)	(0.39)	(0.40)	(0.45)
`treat_wage`	0.34	0.27	0.43	-0.25	-0.38	-0.58	0.37	0.39	0.59
	(0.54)	(0.54)	(0.62)	(0.34)	(0.36)	(0.43)	(0.40)	(0.40)	(0.46)
`treat_cpi` $\times$ $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.03	-0.02	-0.05	-0.07***	-0.05*	-0.01	-0.03	-0.04	-0.06
	(0.06)	(0.06)	(0.07)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.06)
`treat_wage` $\times$ $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.13**	-0.13**	-0.13**	-0.02	0.00	-0.07*	-0.08	-0.08	-0.14**
	(0.05)	(0.05)	(0.06)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)	(0.06)
$\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.51***	-0.52***	-0.46***	-0.88***	-0.89***	-0.92***	-0.44***	-0.48***	-0.44***
	(0.04)	(0.04)	(0.05)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.05)
Constant	2.29***	-1.46	-0.17	2.21***	2.55	2.87	3.01***	3.29***	3.36***
	(0.43)	(2.40)	(2.78)	(0.27)	(1.66)	(1.96)	(0.30)	(1.02)	(1.11)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	1406	1378	1068	1406	1378	1068	1406	1378	1068
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Part 2: Dependent variable: $\Delta\mathbb{E}_{it}^{\texttt{prior3-prior1}}[Z_{t+12}]$									
`treat_cpi`	0.80	0.65	1.12*	0.28	0.03	0.26	-0.21	-0.27	0.01
	(0.53)	(0.54)	(0.63)	(0.32)	(0.33)	(0.39)	(0.41)	(0.42)	(0.47)
`treat_wage`	0.48	0.39	0.79	0.03	-0.09	-0.04	-1.10***	-0.96**	-0.81
	(0.51)	(0.52)	(0.61)	(0.31)	(0.32)	(0.38)	(0.42)	(0.43)	(0.50)
`treat_cpi` $\times$ $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.12**	-0.11**	-0.12*	-0.10***	-0.07***	-0.07**	0.05	0.07	0.03
	(0.05)	(0.05)	(0.06)	(0.02)	(0.02)	(0.03)	(0.05)	(0.05)	(0.06)
`treat_wage` $\times$ $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.18***	-0.18***	-0.16***	-0.06**	-0.04	-0.05	0.19***	0.17***	0.15**
	(0.05)	(0.05)	(0.06)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.07)
$\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.57***	-0.57***	-0.55***	-0.86***	-0.88***	-0.88***	-0.65***	-0.69***	-0.63***
	(0.04)	(0.04)	(0.05)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.05)
Constant	2.02***	1.54	1.11	2.06***	6.59***	8.61***	4.06***	5.33***	4.85***
	(0.41)	(2.31)	(2.70)	(0.25)	(1.48)	(1.74)	(0.32)	(1.09)	(1.19)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	1425	1389	1064	1425	1389	1064	1425	1389	1064
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output from equation (C.2) for $j = 1, 2$. For each outcome variable specified in the header, the first column reports results without controls, the second column adds control variables, and the third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.3.2 Learning Through Survey Effects

This section examines the treatment effects of information provision on expectations in the follow-up waves. Because respondents in the treatment groups have learned about either current CPI inflation rates or current hourly earnings inflation rates by participating in the first wave of the survey, the information treatment effect from subsequent follow-up surveys might be weaker. We explore the possibility of having this "learning-through-survey" effect in this section.

To that end, we run the following regression:

$$\begin{aligned}\Delta\mathbb{E}_{it}^{\text{post-prior}}[Z_{t+12}] =& \beta_0 + \mathbb{1}_{\texttt{wave1}} \times \left(\beta_1^{\texttt{wave1}}\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}] + \sum_{k=1}^{2}\beta_{2,k}^{\texttt{wave1}}\texttt{treat}_i^k + \sum_{k=1}^{2}\beta_{3,k}^{\texttt{wave1}}\left(\texttt{treat}_i^k \times \mathbb{E}_{it}^{\text{prior}}[Z_{t+12}]\right)\right)\\ &+ \mathbb{1}_{\texttt{wave2}} \times \left(\beta_1^{\texttt{wave2}}\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}] + \sum_{k=1}^{2}\beta_{2,k}^{\texttt{wave2}}\texttt{treat}_i^k + \sum_{k=1}^{2}\beta_{3,k}^{\texttt{wave2}}\left(\texttt{treat}_i^k \times \mathbb{E}_{it}^{\text{prior}}[Z_{t+12}]\right)\right)\\ &+ \mathbb{1}_{\texttt{wave3}} \times \left(\beta_1^{\texttt{wave3}}\mathbb{E}_{it}^{\text{prior}}[Z_{t+12}] + \sum_{k=1}^{2}\beta_{2,k}^{\texttt{wave3}}\texttt{treat}_i^k + \sum_{k=1}^{2}\beta_{3,k}^{\texttt{wave3}}\left(\texttt{treat}_i^k \times \mathbb{E}_{it}^{\text{prior}}[Z_{t+12}]\right)\right) + \varepsilon_i,\end{aligned} \tag{C.3}$$

for $Z = \{\pi, \pi^2, u\}$ with those who participated in all three waves of the surveys (937 out of 2,763).[1] By comparing the regression coefficients on the interaction terms between the treatment dummies with prior expectations across three waves ($\beta_{3,k}^{\texttt{wave1}} - \beta_3^{\texttt{wave3}}$), we examine if participants learn through surveys.

Table C.7 shows the estimation results from equation (C.3). First, columns 1-3 in Table C.7 show clear treatment effects of information provisions on expected price inflation rates in the subsequent waves.[2] When respondents receive information about either current CPI inflation rates or hourly earnings inflation rates, they revise their expectations about price inflation rates significantly by putting smaller weights on their priors. The information treatment effects with CPI inflation treatment are of similar magnitudes between the first and the second waves but they become much smaller in the third wave. In contrast, the information treatment effects with hourly earnings treatment are similar between the first and the third waves and they are imprecisely estimated in the second wave.

[1]We followed up with participants in the two treatment groups (CPI and hourly earnings group) and one control group (air quality index group) in the second and third waves. Among 3,979 participants in the first wave, 2,763 of them are in these groups.

[2]See Appendix C.3.3 the estimation results with the full sample who participated in either wave 2 or wave 3.

Table C.7: Effects of information treatments on revision of price inflation, wage inflation, and unemployment expectations (Wave 1-3)

Dependent variable:	Price inflation ($Z = \pi$)			Wage inflation ($Z = \pi^2$)			Unemployment rate ($Z = u$)		
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Wave 1 ×									
`treat_cpi`	1.55***	1.55***	1.59***	1.55***	1.39***	1.33***	-0.74***	-0.76***	-0.48**
	(0.29)	(0.29)	(0.28)	(0.24)	(0.25)	(0.23)	(0.23)	(0.24)	(0.23)
`treat_wage`	1.11***	1.06***	0.95***	2.82***	2.76***	3.63***	-0.04	-0.11	-0.11
	(0.26)	(0.27)	(0.26)	(0.23)	(0.24)	(0.22)	(0.24)	(0.25)	(0.25)
`treat_cpi` × $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.24***	-0.24***	-0.24***	-0.13***	-0.09***	-0.41***	0.13***	0.14***	0.07**
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
`treat_wage` × $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.12***	-0.11***	-0.09***	-0.17***	-0.15***	-0.73***	-0.01	0.00	-0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)
Wave 2 ×									
`treat_cpi`	1.78***	1.74***	1.92***	0.74***	0.59**	1.16***	-0.43*	-0.55**	-0.48**
	(0.28)	(0.28)	(0.27)	(0.23)	(0.24)	(0.22)	(0.24)	(0.25)	(0.23)
`treat_wage`	0.43*	0.36	0.43*	2.28***	2.18***	2.79***	0.86***	0.87***	0.31
	(0.26)	(0.26)	(0.26)	(0.22)	(0.23)	(0.21)	(0.24)	(0.25)	(0.24)
`treat_cpi` × $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.25***	-0.25***	-0.29***	-0.29***	-0.28***	-0.37***	0.09***	0.10***	0.10***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
`treat_wage` × $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.04	-0.03	-0.05	-0.51***	-0.51***	-0.55***	-0.17***	-0.17***	-0.06*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Wave 3 ×									
`treat_cpi`	1.34***	1.21***	1.12***	0.54**	0.11	0.39*	-0.19	-0.22	-0.48**
	(0.27)	(0.28)	(0.27)	(0.23)	(0.23)	(0.22)	(0.24)	(0.25)	(0.24)
`treat_wage`	1.08***	0.99***	1.10***	1.81***	1.77***	2.66***	-0.11	-0.20	0.20
	(0.26)	(0.26)	(0.25)	(0.22)	(0.23)	(0.21)	(0.24)	(0.25)	(0.23)
`treat_cpi` × $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.08**	-0.07**	-0.07*	-0.27***	-0.13***	-0.06**	0.06*	0.07**	0.11***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
`treat_wage` × $\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.16***	-0.16***	-0.13***	-0.35***	-0.36***	-0.52***	0.01	0.03	-0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	2808	2744	2293	2811	2747	2295	2811	2747	2295
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output from equation (C.3). For each outcome variable specified in the header, the first column reports results without controls, the second column adds control variables, and the third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We can observe similar patterns from columns 4-6 in Table C.7. While we observe clear information treatment effects from the follow-up surveys, the information treatment effects become smaller in the third wave. That is, at least for highly numerate respondents, the information treatment effects of CPI treatment and/or hourly earnings treatment become smaller in the third wave as they learn through participating in surveys.

Meanwhile, we do not observe such a pattern from columns 7-9 in Table C.7. Across all waves, respondents further corroborated their priors on unemployment expectations when they received CPI inflation signals. The regression coefficients on the interaction terms between CPI treatment and prior unemployment expectations are statistically significantly positive across all three waves and they are of similar magnitudes. The information treatment effects of hourly earnings treatment on unemployment expectations, on the other hand, are only significant and negative in the second wave. They are imprecisely estimated in the first and the third waves.

C.3.3 Information Treatment Effects From Wave 2 & Wave 3

Lastly, we present the treatment effects of information provision from the second and the third waves of the survey with full observations including those who have participated in either wave 1 and wave 2 or wave 1 and wave 3 only. Table C.8 and C.9 show the results. Consistent with the results in section C.3, they show clear information treatment effects. At the same time, however, the information treatment effects of CPI inflation rates become smaller for the price inflation and unemployment expectations in the third wave. In contrast, the information treatment effects on hourly earnings inflation expectations are of similar magnitudes across the three waves across various treatments.

Table C.8: Effects of information treatments and prior expectations on forecasts revisions from the Wave 2

Dependent variable:	Price inflation ($Z=\pi$)			Wage inflation ($Z=\pi^w$)			Unemployment rate ($Z=u$)		
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
`treat_cpi`	2.14***	2.13***	2.47***	1.28***	1.27***	1.49***	-0.19	-0.20	-0.32
	(0.32)	(0.33)	(0.33)	(0.27)	(0.28)	(0.27)	(0.31)	(0.33)	(0.32)
`treat_wage`	0.84***	0.76**	0.88***	2.00***	2.25***	2.61***	0.71**	0.84**	0.60*
	(0.31)	(0.33)	(0.32)	(0.27)	(0.28)	(0.26)	(0.31)	(0.34)	(0.32)
`treat_cpi` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.34***	-0.34***	-0.36***	-0.28***	-0.33***	-0.45***	0.05	0.06	0.08**
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
`treat_wage` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.14***	-0.14***	-0.14***	-0.32***	-0.43***	-0.54***	-0.14***	-0.15***	-0.11***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
$\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.19***	-0.19***	-0.15***	-0.49***	-0.39***	-0.23***	-0.14***	-0.13***	-0.14***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	1461	1398	1193	1464	1400	1195	1464	1400	1195
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output for respondents who participated in the second wave of the survey from equation (3.2). For each outcome variable specified in the header, the first column reports results without controls, the second column adds control variables, and third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates. Standard errors in parentheses: $^{*}\ p<0.10$, $^{**}\ p<0.05$, $^{***}\ p<0.01$.

Table C.9: Effects of information treatments and prior expectations on forecasts revisions from the Wave 3

Dependent variable:	Price inflation ($Z=\pi$)			Wage inflation ($Z=\pi^w$)			Unemployment rate ($Z=u$)		
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
`treat_cpi`	1.57***	1.50***	1.52***	1.06***	0.77***	0.79***	0.15	0.12	-0.08
	(0.32)	(0.32)	(0.31)	(0.23)	(0.24)	(0.23)	(0.30)	(0.33)	(0.31)
`treat_wage`	0.77**	0.73**	1.10***	2.11***	2.09***	2.50***	-0.04	-0.22	0.11
	(0.31)	(0.31)	(0.30)	(0.23)	(0.24)	(0.23)	(0.30)	(0.32)	(0.31)
`treat_cpi` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.18***	-0.19***	-0.16***	-0.39***	-0.33***	-0.28***	-0.02	-0.01	0.03
	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
`treat_wage` $\times\ \mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.18***	-0.17***	-0.18***	-0.43***	-0.43***	-0.51***	-0.01	0.01	-0.04
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
$\mathbb{E}_{it}^{\texttt{prior}}[Z_{t+12}]$	-0.27***	-0.26***	-0.17***	-0.29***	-0.34***	-0.26***	-0.07**	-0.09***	-0.10***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Sample	All	All	Numerate	All	All	Numerate	All	All	Numerate
N	1468	1432	1223	1470	1434	1225	1470	1434	1225
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: This table presents the Huber-Robust regression output for respondents who participated in the second wave of the survey from equation (3.2). For each outcome variable specified in the header, the first column reports results without controls, the second column adds control variables, and third column restricts the sample to highly numerate respondents only (who answered all the numerical competence questions correctly). Control variables are female, age, age^2, white, whether cohabiting or not, whether having a child or not, full-time employed or not, logarithmic monthly spending on food, hours working at MTurk, whether having a college degree or not, frequency of checking news, and income). The control group refers to those who have received irrelevant information such as the air quality index in Seattle or Covid-19 vaccination rates. Standard errors in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

C.4 Robustness Checks

C.4.1 Adjusted of $p-$values for Multiple Hypothesis Testing

First, we provide the same regression results with the adjusted p-values for multiple hypothesis testing. In the second stage, we are interested in six parameter values. The regression coefficients on the forecast revisions in price and wage inflation rates, and unemployment rates with two dependent variables: the desired duration of employment and the reservation wages. To minimize the likelihood of false rejections with multiple hypothesis testing, we use Westfall-Young stepdown adjusted p-values using `wyoung` command in STATA. This controls the familywise error rate (FWER) and allow for dependence amongst p-values. The results are reported in Table C.10

Results in Panel A of Table C.10 replicate the results in Sections 3.4.1 and 3.4.2. They are similar to the baseline results. In terms of desired duration of employment on MTurk project, the results point to a statistically significant and positive effect of unemployment rate expectations on employment duration, and negative effect of wage growth expectations on employment duration. As to the reservation wages, a positive effect of wage inflation expectations and negative effect of price inflation expectations are largely robust to standard errors adjustment. Results in Panel B of Table C.10 also closely matches the results in Table 3.9 about broad regime changes.

Table C.10: Effects of expectations on MTurk labor supply with adjusted p-values

	Desired Duration (in months)				Reservation Wages			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Continuous priors								
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[\pi_{t+12}]$	-0.12	-1.26	0.56	0.22	-0.50	-0.32	-1.38	-1.47*
	(0.98)	(0.78)	(0.78)	(0.78)	(0.86)	(0.76)	(0.22)	(0.10)
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[\pi_{t+12}^{w}]$	-0.86	0.01	-2.10**	-1.70*	2.22*	2.88*	1.20***	0.73
	(0.98)	(0.99)	(0.01)	(0.06)	(0.06)	(0.03)	(0.01)	(0.23)
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[u_{t+12}]$	4.48**	3.12	4.41***	2.72*	-1.72*	0.17	0.26	0.93
	(0.02)	(0.23)	(0.00)	(0.06)	(0.07)	(1.00)	(0.78)	(0.13)
N	3,141	3,079	2,222	2,160	3,075	3,015	2,110	2,056
Sample	All	All	Numerate	Numerate	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}]$	11.87	10.42	22.46	20.85	10.84	10.51	15.68	15.01
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}^{w}]$	12.18	10.73	50.03	40.53	12.71	11.61	54.64	37.89
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[u_{t+12}]$	31.04	29.86	49.91	52.05	36.80	31.37	48.21	45.64
Panel B: Broad regime changes								
$\texttt{Regime}^{\pi}$	-0.03	0.03	-0.00	0.10	-0.13**	-0.02	-0.04	-0.08
	(0.84)	(0.94)	(0.97)	(0.90)	(0.03)	(0.62)	(0.55)	(0.44)
$\texttt{Regime}^{\pi^{w}}$	-0.17	-0.25**	-0.22*	-0.24**	0.17***	0.20***	0.12*	0.19**
	(0.14)	(0.02)	(0.09)	(0.04)	(0.01)	(0.00)	(0.09)	(0.02)
$\texttt{Regime}^{u}$	0.47***	0.46***	0.41***	0.35**	-0.14**	-0.03	-0.04	0.04
	(0.00)	(0.00)	(0.00)	(0.02)	(0.04)	(0.94)	(0.76)	(0.90)
N	3,127	3,088	2,203	2,158	3,051	3,005	2,118	2,080
Sample	All	All	Numerate	Numerate	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-stat for $\texttt{Regime}^{\pi}$	29.73	26.94	16.81	15.34	26.70	22.47	13.71	10.43
F-stat for $\texttt{Regime}^{\pi^{w}}$	36.40	34.60	24.69	27.38	32.33	34.31	21.38	22.49
F-stat for $\texttt{Regime}^{u}$	41.77	30.43	40.57	30.67	36.08	24.92	37.25	26.94

Notes: This table presents the regression output to estimate the effects of expectations on MTurk labor supply with the *adjusted* p-values in parentheses. To minimize the likelihood of false rejections with multiple hypothesis testing, we use Westfall-Young stepdown adjusted p-values using `wyoung` command in STATA. Panel A shows the results from equation (3.3) and Panel B shows the results from equation (3.5). We instrument the revisions in expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, and the interaction of unemployment treatment dummies with prior expected unemployment rates. Highly numerate respondents are those who answered all the numerical competence check questions correctly. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for details about the treatment of outliers.

C.4.2 Alternative Instruments

Next, we provide the estimation results with a different set of instrumental variables. In addition to the instruments we have in Section 3.4, we add the interaction of $\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}]$ with the treatment dummy for unemployment rates, the interaction of $\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}^{w}]$ with the treatment dummy for unemployment rates, and the interactions of $\mathbb{E}_{it}^{\texttt{post}}[u_{t+12}]$ with the treatment dummies for CPI and hourly earnings inflation rates to the set of instruments. That is, our full set of instruments are now: $\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[Z_{t+12}]$ for $Z \in \{\pi, \pi^2, u\}$ with the following set of IVs: $\texttt{treat\_cpi}_{it}$, $\texttt{treat\_wage}_{it}$, $\texttt{treat\_unemp}_{it}$, $(\texttt{treat\_cpi}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}])$, $(\texttt{treat\_cpi}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}^{w}])$, $(\texttt{treat\_wage}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}^{w}])$, $(\texttt{treat\_wage}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}])$, $(\texttt{treat\_unemp}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[u_{t+12}])$, $(\texttt{treat\_unemp}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}])$, $(\texttt{treat\_unemp}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[\pi_{t+12}^{w}])$, $(\texttt{treat\_cpi}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[u_{t+12}])$, and $(\texttt{treat\_wage}_{it} \times \mathbb{E}_{it}^{\texttt{prior}}[u_{t+12}])$.

Table C.11 shows the regression results from the same regression models of equations (3.3) and (3.5) with these instrumental variables. As can be seen from Panel A of Table C.11, the results are consistent with those in the main text with smaller Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests. Higher unemployment expectations are associated with a higher desired duration of employment. Higher wage inflation expectations increase reservation wages. In contrast, higher price inflation expectations rather decrease reservation wages for highly numerate respondents.

Panel B of Table **??** shows that the results are consistent with the baseline results for broad regime changes in Section 3.5.2. Broad changes in the price inflation regime do not affect the desired duration of employment, but the broad changes in hourly earnings inflation do. As respondents move from no change to an increase in expected hourly earnings growth rate or from a decrease no change or increase in expected hourly earnings growth rate, they decrease their desired duration of employment. As respondents increase their unemployment expectation *upwards*, they increase their desired duration of employment. Moreover, the last four columns show that as respondents revise their broad regime about hourly earnings inflation expectation *upwards*, they increase their reservation wages. In contrast, the upward revision of price inflation expectations is associated with the decrease in reservation wages. Similarly, the upward forecast revisions of unemployment rates are associated with lower reservation wages.

Table C.11: Effects of expectations on MTurk labor supply with additional instruments

	Desired Duration (in months)				Reservation Wages			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Continuous priors								
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[\pi_{t+12}]$	-0.42	-0.86	1.10	0.84	-0.54	-0.79	-1.46**	-1.27*
	(1.20)	(1.20)	(0.97)	(0.96)	(0.73)	(0.69)	(0.65)	(0.66)
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[\pi_{t+12}^{w}]$	-1.29	-0.79	-1.27	-2.57**	2.72***	3.09***	1.83***	1.35***
	(1.85)	(1.70)	(1.14)	(1.30)	(0.93)	(0.89)	(0.47)	(0.48)
$\Delta\mathbb{E}_{it}^{\texttt{post-prior}}[u_{t+12}]$	3.81***	3.10**	3.72***	3.94***	-1.70*	0.18	-0.29	0.82
	(1.46)	(1.43)	(1.14)	(1.16)	(0.87)	(0.91)	(0.72)	(0.81)
N	3,155	3,085	2,248	2,210	3,068	3,009	2,134	2,070
Sample	All	All	Numerate	Numerate	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}]$	7.73	7.00	12.21	10.62	7.32	8.00	9.27	9.50
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}^{w}]$	8.42	7.56	19.40	11.75	8.36	8.21	18.13	19.59
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[u_{t+12}]$	19.71	18.52	31.98	30.21	21.08	16.33	33.54	21.66
Panel B: Broad regime changes								
$\texttt{Regime}^{\pi}$	-0.05	0.05	-0.03	0.05	-0.13**	-0.05	-0.13*	-0.06
	(0.09)	(0.10)	(0.11)	(0.12)	(0.06)	(0.06)	(0.07)	(0.08)
$\texttt{Regime}^{\pi^{w}}$	-0.17*	-0.17*	-0.22**	-0.26***	0.23***	0.19***	0.13**	0.11*
	(0.09)	(0.10)	(0.09)	(0.10)	(0.05)	(0.05)	(0.05)	(0.06)
$\texttt{Regime}^{u}$	0.27***	0.32***	0.29***	0.39***	-0.21***	-0.13**	-0.11**	-0.07
	(0.10)	(0.11)	(0.09)	(0.10)	(0.06)	(0.06)	(0.05)	(0.06)
N	3,146	3,108	2,246	2,190	3,108	3,049	2,180	2,118
Sample	All	All	Numerate	Numerate	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-stat for $\texttt{Regime}^{\pi}$	26.75	25.04	13.91	12.52	25.34	22.76	13.17	12.16
F-stat for $\texttt{Regime}^{\pi^{w}}$	28.14	27.23	19.05	17.52	27.84	28.49	17.99	16.59
F-stat for $\texttt{Regime}^{u}$	18.33	14.24	21.21	16.82	17.04	14.63	21.40	16.72

This table presents the regression output to estimate the effects of expectations on MTurk labor supply. Panel A shows the results from equation (3.3) and Panel B shows the results from equation (3.5). We instrument the revisions in expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, the interaction of unemployment treatment dummies with prior expected unemployment rates, the interaction of prior price inflation expectations with the treatment dummy for unemployment rates, the interaction of prior wage inflation expectations with the treatment dummy for unemployment rates, and the interactions of prior expected unemployment rates with the treatment dummies for CPI and hourly earnings inflation rates. Heteroskedasticity-robust-standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Highly numerate respondents are those who answered all the numerical competence check questions correctly. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for the treatment of outliers.

C.4.3 Alternative Dependent Variable

Finally, we use the alternative regression specification to estimate the effect of macroeconomic expectations on MTurk labor supply in levels rather than changes. Specifically, we estimate the following regression:

$$\begin{aligned} \mathtt{Y}_{it}^{\mathtt{post}} =& \beta_0 + \beta_1 \Delta \mathbb{E}_{it}^{\mathtt{post\text{-}prior}}[\pi_{t+12}] + \beta_2 \Delta \mathbb{E}_{it}^{\mathtt{post\text{-}prior}}[\pi_{t+12}^{w}] + \beta_3 \Delta \mathbb{E}_{it}^{\mathtt{post\text{-}prior}}[u_{t+12}] \\ &+ \gamma_1 \mathbb{E}_{it}^{\mathrm{prior}}[\pi_{t+12}] + \gamma_2 \mathbb{E}_{it}^{\mathrm{prior}}[\pi_{t+12}^{w}] + \gamma_3 \mathbb{E}_{it}^{\mathrm{prior}}[u_{t+12}] + \gamma_4 \mathtt{Y}_{it}^{\mathtt{prior}} \mathbf{X}_{it}' \delta + \varepsilon_i, \end{aligned} \quad \text{(C.4)}$$

where $Y_{it} = \{\mathtt{dur}_{it}^{\mathtt{post}}, \overline{\mathtt{rw}}_{it,t+\mathtt{dur}_t}^{\mathtt{post}}\}$ are desired duration of employment on our MTurk project (in month) and reservation wage per 10-minute monthly task. $\mathtt{Regime\ Change}_i^Z$ is an indicator variable denoting if respondent i revises her *qualitative* assessment about a variable Z *upwards* defined in the same way is in Section 3.5.1.

Table C.12 shows the results that are broadly consistent with the results from Tables 3.4 and 3.5 in Section 3.4. The main difference here is that now the desired duration of employment does not significantly increase with higher expected unemployment rates. The results for the reservation wages are, however, consistent with the results in the main text. Higher wage inflation expectations are associated with higher reservation wages. A one percentage point increase in expected wage inflation rates increases reservation wages by 1-2 cents. Moreover, for highly numerate respondents, reservation wage decreases with higher price inflation expectations, reflecting the stagflationary view of the U.S. households.

Table C.12: Effects of expectations on MTurk labor supply in labor supply in levels

	Desired Duration (in months)				Reservation Wages			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{E}_{it}^{\texttt{post-prior}}[\pi_{t+12}]$	-0.17	-0.09	-0.12	-0.06	0.09	-0.78	-1.48**	-0.67
	(0.28)	(0.22)	(0.19)	(0.23)	(0.68)	(0.62)	(0.61)	(0.61)
$\mathbb{E}_{it}^{\texttt{post-prior}}[\pi^{w}_{t+12}]$	0.00	-0.44	0.01	-0.44	1.68**	0.88**	1.01***	0.49
	(0.44)	(0.27)	(0.16)	(0.29)	(0.79)	(0.40)	(0.37)	(0.36)
$\mathbb{E}_{it}^{\texttt{post-prior}}[u_{t+12}]$	-0.47	0.04	-0.04	-0.02	-1.45*	0.66	0.62	0.13
	(0.29)	(0.24)	(0.18)	(0.23)	(0.80)	(0.72)	(0.73)	(0.64)
N	3,112	2,173	2,192	2,154	2,992	2,076	2,084	2,008
Sample	All	All	Numerate	Numerate	All	All	Numerate	Numerate
Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi_{t+12}]$	13.61	15.12	20.45	18.51	10.99	13.14	13.89	13.06
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[\pi^{w}_{t+12}]$	13.35	19.49	51.61	33.82	13.58	28.15	33.54	31.02
F-stat for $\Delta\mathbb{E}_{it}^{\texttt{post}}[u_{t+12}]$	31.08	43.20	45.98	41.29	31.51	36.41	44.23	49.23

Notes: This table presents the regression output to estimate the effects of expectations on MTurk labor supply according to equation (3.3). The first four columns show the Poisson regression results. We instrument the posterior expectations with the treatment dummies of CPI inflation rates, hourly earnings inflation rates, and unemployment rates, the interactions of prior price inflation expectations with the CPI inflation treatment dummies and with the hourly earnings inflation treatment dummies, the interactions of prior wage inflation expectations with the CPI inflation treatment dummies and with the hourly earnings treatment dummies, and the interaction of unemployment treatment dummies with prior expected unemployment rates. Heteroskedasticity-robust-standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Highly numerate respondents are those who answered all the numerical competence check questions correctly. Kleibergen and Paap (2006) rk Wald F-statistics for weak identification tests are reported. We use the geometric average of the weights generated from the Huber-robust regressions for each variable of interest in the first stage to control for outliers of the variables regarding expectations. To control for outliers in the second stage, we use a jackknife approach. See Appendix C.5 for the details about treatment of outliers.

C.5 Treatment of Outliers

To deal with outliers in expectations and labor supply data, we use the strategy following Coibion et al. (2023). To be more specific, we use the Huber-robust regression in the first stage with `rreg` command in STATA.[3] In this process, we generate weights to deal with outliers in the subjective expectations data. We run the second stage using the weights generated from the first stage. Because we run three first-stage regressions with posterior price, wage inflation expectations, and expected unemployment rates, we have three weights generated from the first stage. We take the geometric average over the three weights and use it in the second stage.

To further remove the influence of outliers in the second stage, we use the jackknife approach in the second stage. That is, we calculate the regression coefficients by dropping one observation each to find influential observations. We then drop observations as long as it moves the regression coefficients on posterior expectations by a magnitude greater than 0.07.[4]

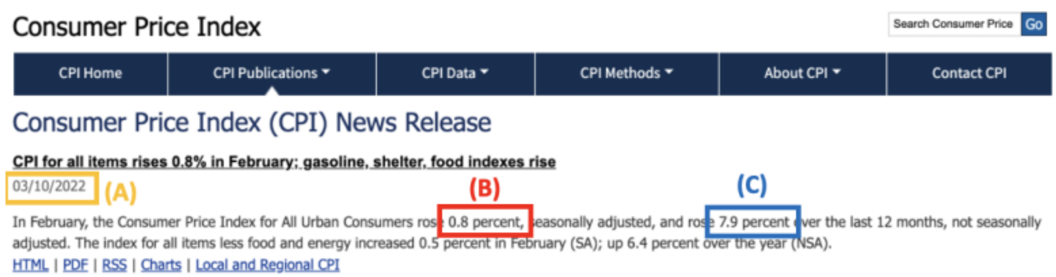
[3]For more detail, see help for STATA's `rreg` command. Or see Appendix C of Coibion et al. (2023).

[4]Besley et al. (1980) suggests to use the threshold of $2/\sqrt{N}$, where N is the number of observations. After dropping the duplicated observations, we have 3,979 observations in the first wave. This corresponds to the threshold of 0.0317. We pick a higher number to drop a smaller number of observations. Our results are robust to the choice of this value from 0.05 to 0.10.

C.6 Examples of the Main Task

Treatment groups

Based on the information from this screenshot, please fill the table below it.



Consumer Price Index

Search Consumer Price Go

CPI Home | CPI Publications | CPI Data | CPI Methods | About CPI | Contact CPI

Consumer Price Index (CPI) News Release

CPI for all items rises 0.8% in February; gasoline, shelter, food indexes rise

03/10/2022 (A)

(B) (C)

In February, the Consumer Price Index for All Urban Consumers rose 0.8 percent, seasonally adjusted, and rose 7.9 percent over the last 12 months, not seasonally adjusted. The index for all items less food and energy increased 0.5 percent in February (SA); up 6.4 percent over the year (NSA).

HTML | PDF | RSS | Charts | Local and Regional CPI

Source: https://www.bls.gov/cpi/news.htm

Table

	Date of the news report	CPI inflation rate	
	mm/dd/yyyy (A)	in March 2022, in percent (B)	over the last 12 months, in percent (C)
Your answer			

Figure C.3: Example of text transcription task: CPI inflation rate

D. Based on the information from this screenshot, please fill the table below it.

Table B-3. Average hourly and weekly earnings of all employees on private nonfarm p

ESTABLISHMENT DATA
Table B-3. Average hourly and weekly earnings of all employees on private nonfarm payrolls by industry sector, seasona

Industry	Average hourly earnings			
	Mar. 2021 (B)	Jan. 2022	Feb. 2022(P)	Mar. 2022(P) (C)
Total private	$30.06	$31.56	$31.60	$31.73
Goods-producing	30.45	31.91	31.88	31.97
Mining and logging	34.30	35.90	35.75	35.75
Construction	32.24	33.87	33.94	34.07
Manufacturing	29.20	30.57	30.46	30.55
Private service-providing	29.97	31.48	31.54	31.67
Trade, transportation, and utilities	25.83	27.14	27.26	27.44
Information	44.06	44.77	45.18	45.17
Financial activities	39.77	40.88	40.87	41.19
Professional and business services	35.80	37.92	37.97	38.18
Education and health services	29.46	31.22	31.25	31.24
Leisure and hospitality	17.60	19.43	19.45	19.68
Other services	27.22	28.37	28.31	28.18

Footnotes
(P) Preliminary

Last Modified Date April 01, 2022 (A)

Source: https://www.bls.gov/news.release/empsit.t19.htm#ces_table3.f.p

Table.

	Date when table was last modified	Average hourly earnings of all employees in the private sector in the U.S. (omit $ symbol)	
	mm/dd/yyyy (A)	in March **2021** (B)	in March **2022** (C)
Your answer			

Figure C.4: Example of text transcription task: Hourly earnings

Based on the information from this screenshot, please fill the table below it.

U.S. Unemployment Rate Forecast

U.S. Unemployment Rate Forecast Values
Percent Unemployed, Seasonally Adjusted.

Month	Date	Forecast Value	Avg Error
0	Mar 2022	(B) 3.6	±0.0
1	Apr 2022	3.6	±0.08
2	May 2022	3.5	±0.1
3	Jun 2022	3.5	±0.1
4	Jul 2022	3.5	±0.1
5	Aug 2022	3.5 (C)	±0.1
6	Sep 2022	3.4	±0.2
7 (A)	Oct 2022	3.4	±0.2
8	Nov 2022	3.4	±0.2

Modified: April 04, 2022

Source: https://www.forecasts.org/unemploy.htm

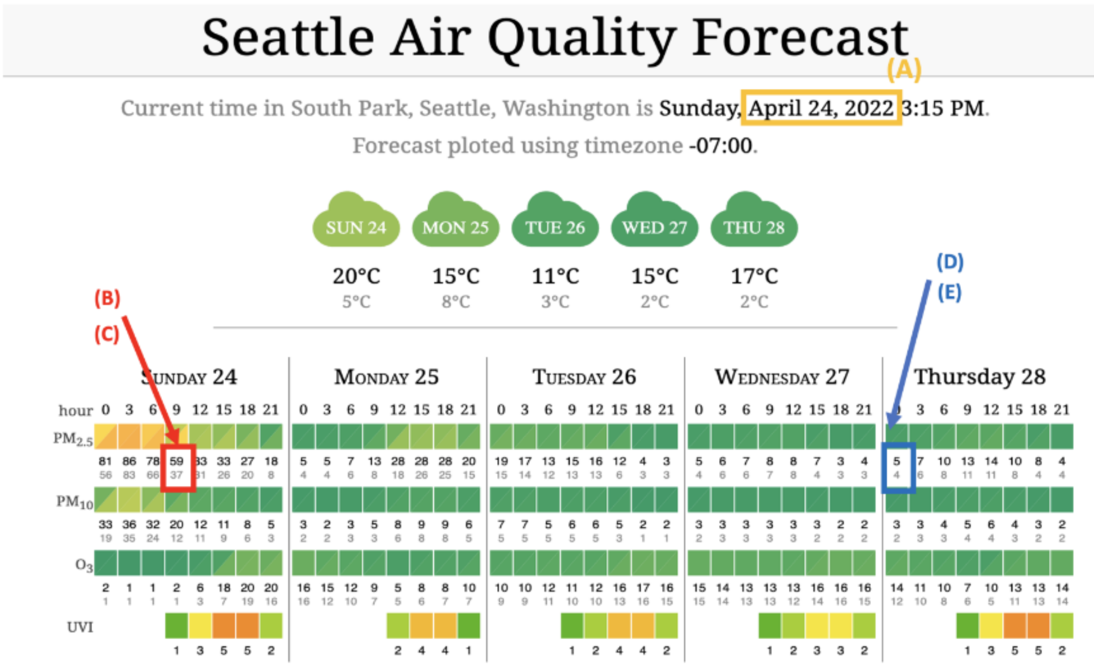
Table.

	Date when the table is last modified mm/dd/yyyy **(A)**	Unemployment rate in the **previous month**, in percent **(B)**	Unemployment rate forecast in **six months**, in percent **(C)**
Your answer			

Figure C.5: Example of text transcription task: Unemployment rate

Control groups

Based on the information from this screenshot, please fill the table below it.



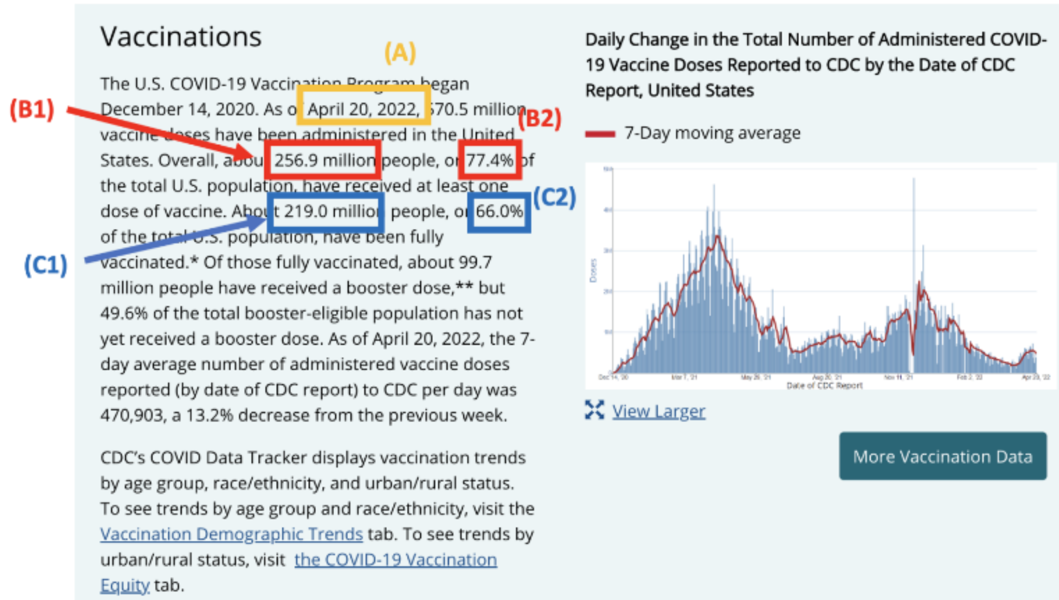
Source: https://aqicn.org/forecast/seattle/

Table.

	Date of the forecast	What is the air quality index (PM 2.5) at 12 pm on the day of the forecast?		What is the forecast for the air quality index (PM 2.5) at 12 pm in 4 days?	
	mm/dd/yyyy (A)	High (B)	Low (C)	High (D)	Low (E)
Your answer					

Figure C.6: Example of text transcription task: Air quality index

D. Based on the information from this screenshot, please fill the table below it.



Source: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html

Table.

	Date of the report	Total COVID Vaccination Rates (At least one dose)		Total COVID Vaccination Rates (Fully Vaccniated)	
	mm/dd/yyyy **(A)**	Count (in millions) **(B1)**	Percent of the US Population **(B2)**	Count (in millions) **(C1)**	Percent of the US Population **(C2)**
Your answer					

Note. Ignore % symbol during transcription.

Figure C.7: Example of text transcription task: Covid-19 vaccination rate

C.7 Survey Questions

Start of Block: Description

Consent

CONSENT TO PARTICIPATE IN HIT "SHORT SURVEY + FORECASTING TASK"

Please find below information about this HIT for you to carefully consider when deciding about whether to participate. Please ask questions about any of the information you do not understand before you decide whether to participate.

Contact Information:
EpiLS Study Team
Email: Epilsstudyteam@tufts.edu
Phone: 617-627-3560

We are collecting data for training a machine learning forecasting model. Once our study is completed, we will provide you with full information.

In this task, in addition to answering several questions about you and your experience, we ask you to:
1) Transcribe the statistical information from a screenshot
2) Record your own forecasts based on the information provided.

Before the main task, you will be asked to do a short screening task on transcribing text from a screenshot. Only after you complete the screening task accurately, you will be eligible to proceed with the remainder of the study.

It takes about 10-15 minutes to complete this HIT.

Once your HIT is approved, you will be paid $1.50.

HIT approval decision will be based on the following three criteria: i) survey completion, ii) accuracy of transcription, and iii) quality of your answers. If your answers are meaningful, you transcribe the information accurately, and you complete the survey, your HIT will be approved.

This HIT includes a few numerical competence checks and transcription of text from a screenshot. They are designed for working on a computer. Some of the tasks might not be mobile-friendly and may cause eye strain.

Participation is completely voluntary. You have the right to quit this HIT at any point. If you quit before completing the survey, however, your HIT will not be approved, and you will not be paid. The data collected to the point of withdrawal will be discarded.

We will take measures to protect your privacy and confidentiality. Although your Mechanical Turk Worker ID will be used to distribute the payment to you, we will not store your worker ID with your survey responses. We will not collect any personally identifiable information except for the encrypted version of your Amazon worker ID. Our research team will have access only to encrypted ID and your anonymized answers which will be stored on password-protected computers. De-identified data will be retained indefinitely for possible use in future research.

Despite taking steps to protect your privacy, we can never fully guarantee your privacy. If you tell us something that makes us believe that you or others have been or may be harmed due to participation in this HIT, we may report that information to the appropriate agencies. Individuals and organizations responsible for conducting or monitoring this study may be permitted to access and inspect the research records. This includes Tufts SBER IRB or Berkeley OPHS.

If you have questions and concerns, contact us. If you go to your Dashboard on MTurk, you can click “Contact Requester” and send us your message.

Institutional Review Boards (“IRB”) are overseeing this study. An IRB is a group of people who perform independent review of studies to ensure the rights and welfare of participants are protected. The research has been approved by IRB boards of the institutions with which researchers are affiliated – Tufts University (STUDY00002463) and University of California, Berkeley (CPHS Protocol 2022-01-14981). If you have questions about your rights or wish to speak with someone other than the research team, you may contact:

Tufts Social, Behavioral, and Educational Research IRB
75 Kneeland Street, Suite 623
Boston, MA 02111
617.627.8804
SBER@tufts.edu

Office for Protection of Human Subjects
University of California, Berkeley
1608 Fourth Street, Suite 220
Mail Code 5940
Berkeley CA, 94710-1749
510-642-7461
ophs@berkeley.edu

STATEMENT OF CONSENT

I have read and considered the information presented in this form. I confirm that I understand the purpose of the study and procedures. I understand that I may ask questions at any time and can withdraw my participation without prejudice. I have read this consent form.

By selecting “I agree,” you are consenting to participate in this study.

- ○ I agree
- ○ I disagree

End of Block: Description

Start of Block: Screening

Screening task Please enter the information from highlighted fields of the screenshot into a table below.

Table 4.2 Real gross domestic product by major demand category, 2000, 2010, 2020, and projected 2030
(Numbers in billions of chained 2012 dollars)

Category	2000	2010	2020 (B)	2030	Compound annual rate of change, 2000–10	Compound annual rate of change, 2010–20 (C)	Compound annual rate of change, 2020–30	Contribution to percent change in real GDP, 2000–10	Contribution to percent change in real GDP, 2010–20
Gross domestic product	$13,131.0	$15,598.7	$18,423.4	$23,029.8	1.7	1.7	2.3	1.7	1.7
Personal consumption expenditures	8,643.3	10,643.0	12,725.9	16,586.0	2.1	1.8	2.7	1.4	1.2
Gross private domestic investment	2,346.7	2,216.5	3,261.2	4,575.5	-0.6	3.9	3.4	-0.1	0.6
Exports	1,379.5	1,977.9	2,216.3	3,171.9	3.7	1.1	3.6	0.4	0.2
Imports(1)	1,930.3	2,543.8	3,142.6	5,098.3	2.8	2.1	5.0	0.4	0.4
Government consumption expenditures and gross investment	2,663.0	3,307.2	3,340.4	3,586.1	2.2	0.1	0.7	0.4	0.0

Footnotes:
(1) Imports are subtracted from the other components of GDP because they are not produced in the United States.
Note: Dash indicates data not computable or not applicable.
Source: Historical data: U.S. Bureau of Economic Analysis; Projected data: U.S. Bureau of Labor Statistics

Top

Last Modified Date September 8, 2021 (A)

Source: https://www.bls.gov/emp/tables/real-gdp-major-demand-category.htm#top

Note: If you transcribe the information incorrectly, you will NOT be permitted to proceed with this HIT.

Table

	Date when table was last modified	Gross Domestic Product in 2020	Compound annual rate of change (2010-20)
	mm/dd/yyyy (A)	in billions USD (B), (ignore all the symbols [e.g. $ and ,] except for decimal points .)	rate (C)
Your answer			

End of Block: Screening

Start of Block: Ba. Prior A - Reservation Wage

B1a The following three questions test your numerical competence.

Anna earns on average $1.00 per 10 minutes of work on MTurk. How much does Anna earn for an hour (60 minutes)?

____

B2a John had earned $8.00 *per hour* before receiving a 5% raise. How much does John earn after the raise *per hour*?

____

B3a A cafe has increased the price of a coffee from $2 to $2.5. How much has the price of a coffee increased *in percent*?

____

B4a Suppose after completing a HIT on MTurk you are offered to participate in a follow-up task. What is the **smallest reward** for a 10-min HIT you would **accept** in **May 2022**? (in USD)

- ○ 0.5
- ○ 0.6
- ○ 0.7
- ○ 0.8
- ○ 0.9
- ○ 1.0
- ○ 1.1
- ○ 1.2
- ○ 1.3
- ○ 1.4
- ○ 1.5
- ○ I would accept a HIT that pays below 0.5 USD
- ○ I would NOT accept any HIT that pays below 1.5 USD

Display This Question:

If Suppose after completing a HIT on MTurk you are offered to participate in a follow-up task. What... = I would accept a HIT that pays below 0.5 USD

Or Suppose after completing a HIT on MTurk you are offered to participate in a follow-up task. What... = I would NOT accept any HIT that pays below 1.5 USD

B5a What is the smallest reward you would accept for a 10-minute HIT?

- ○ Pay for 10 minutes, USD ________________________________

B6a Would you accept work on a HIT that pays **$e{Selected Choice + 0.05} USD** per 10-min session in **May 2022**?

◯ Yes

◯ No

Display This Question:

If the answer to the above question = No

B6a1 What is the **smallest reward** you would **accept** for a 10-minute HIT in **May 2022**?

◯ Pay for 10 minutes, USD

____

B7a
How about a follow-up task that asks you to do a 10-minute HIT **two times -- in May and June 2022**. What is the **smallest reward** for **<u>20 minutes</u>** of your work that you would accept? (in USD)

- ◯ 0.50
- ◯ 0.60
- ◯ 0.70
- ◯ 0.80
- ◯ 0.90
- ◯ 1.00
- ◯ 1.25
- ◯ 1.50
- ◯ 1.75
- ◯ 2.00
- ◯ 2.25
- ◯ 2.50
- ◯ 2.75
- ◯ 3.00
- ◯ 3.25
- ◯ 3.50
- ◯ 3.75
- ◯ 4.00
- ◯ 4.50

○ 5.00

○ 5.50

○ I would accept three HITs that pay less than 0.60 USD for 20 minutes

○ I would NOT accept three HITs that pay less than 5.50 USD for 20 minutes

B8a1 Suppose you could choose for how many months to work on a monthly hit paying ${your answer in B4a or in B5a} USD for 10 minutes of work. For how many **months** would you prefer to work?

○ 0

○ 1

○ 2

○ 3

○ 4

○ 5

End of Block: Bb. Prior A - Reservation Wage

Start of Block: C. Prior - Forecasts

C **FORECASTING TASK**

The next block of questions refers to the main forecasting task. If you are not certain about the answer to any of the following questions, please provide your **best guess**.

Note, we care about **your forecasts**. Therefore, if it is obvious that you have not given any thought to answering the questions and instead entered random numbers, we will not approve your HIT. As long as your answers are meaningful, your HIT will be approved.

To understand what we mean by a *meaningful answer,* see the question below.

C1
Suppose that the question asks "What do you think the average temperature is in Oahu, Hawaii, in **July**? (in Fahrenheit)" and your answer is **30**. Would your HIT be approved?

○ Yes

○ No

C2 What do you think is the average air quality index (AQI) in Seattle, USA was over the **past year**?

○ Mostly good (AQI 0-50)

○ Mostly moderate (AQI 51-100)

○ Unhealthy for sensitive groups (AQI 101-150)

○ Unhealthy (AQI 151-200)

○ Very unhealthy (AQI 201-300)

○ Hazardous (AQI 301-500)

C4 In your opinion, what is the percentage of the U.S. population that has received **at least one dose** of Covid vaccine by today?

____

C5a In each of the scenarios below, what do you think the unemployment rate in the U.S. will be in **April 2023**?

Note: In February 2020, right before the pandemic, the unemployment rate was 3.5%. In April 2020 after the pandemic, the unemployment rate peaked at 14.7%.

- ○ The ***lowest*** possible unemployment rate
  ____
- ○ The ***median*** (or ***average***) unemployment rate
  ____
- ○ The ***highest*** possible unemployment rate
  ____

C5b For each of the scenarios below, please distribute 100 points to indicate how likely you think each unemployment rate will happen. The sum of the points you allocate should total to 100.

The likelihood of the ***lowest*** possible unemployment rate scenario : _______
The likelihood of the ***median*** unemployment rate scenario : _______
The likelihood of the ***highest*** possible unemployment rate scenario : _______
Total : ________

C3a In your opinion, what are the average **hourly** earnings of employees in the private sector in the U.S. in **April 2022**?

◯ Average hourly earnings in April **2022**, USD ________________

C3b In your opinion, will average **hourly** earnings of employees in the U.S. be higher or lower in April **2023** relative to today?

◯ **Higher** than today

◯ About the **same** as today

◯ **Lower** than today

Display This Question:

If In your opinion, will average hourly earnings of employees in the U.S. be higher or lower in Apri... = ***Higher*** *than today*

C3_2a How much **higher** do you think the average hourly earnings in the U.S. will be in April 2023 relative to today (in percentage terms)?

If earnings double over a year, this corresponds to 100% increase. If earnings do not change, this corresponds to 0% increase. E.g., change from 20 to 40 USD corresponds to 100% increase. Change from 20 to 24 USD corresponds to 20% increase. Change from 20 to 21 USD corresponds to 5% increase. Change from 20.0 to 20.2 USD corresponds to 1% increase.

Increase in the average hourly earnings from April 2022 to April 2023:

◯ in percent ________________

Display This Question:

If In your opinion, will average hourly earnings of employees in the U.S. be higher or lower in Apri... = ***Lower*** *than today*

C3_2b How much **lower** do you think the average hourly earnings in the U.S. will be in April 2023 relative to today (in percentage terms)?

If earnings halved over a year, this corresponds to 50% decrease. If earnings do not change, this corresponds to 0% decrease. E.g., change from 20 to 10 USD corresponds to 50% decrease. Change from 20 to 16 USD corresponds to 20% decrease. Change from 20 to 19 USD corresponds to 5% decrease. Change from 20.0 to 19.8 USD corresponds to 1% decrease.

Decrease in the average hourly earnings from April 2022 to April 2023:

○ in percent ____

Display This Question:

*If In your opinion, will average hourly earnings of employees in the U.S. be higher or lower in Apri... = About the **same** as today*

C3_2c You have indicated that you expect that average hourly earnings in the U.S. will be about the same as today in April 2023. This could mean that the change equals zero percent or that the percent change is small. Please select a category that best describes your opinion.

- ○ In April 2023 by 5% lower than today
- ○ In April 2023 by 4% lower than today
- ○ In April 2023 by 3% lower than today
- ○ In April 2023 by 2% lower than today
- ○ In April 2023 by 1% lower than today
- ○ In April 2023 exactly the same as today
- ○ In April 2023 by 1% higher than today
- ○ In April 2023 by 2% higher than today
- ○ In April 2023 by 3% higher than today
- ○ In April 2023 by 4% higher than today
- ○ In April 2023 by 5% higher than today

Page Break

C6a In your opinion, will **prices** in the U.S. be higher or lower in **April 2023** relative to today?

- ◯ **Higher** than today
- ◯ About the **same** as today
- ◯ **Lower** than today

Display This Question:

*If In your opinion, will prices in the U.S. be higher or lower in April 2023 relative to today? = **Higher** than today*

C6a_1 How much do you think the **overall price level** in the U.S. will increase between April **2022** and April **2023** (in percentage terms)?

For example, if cost of a typical consumer basket increases from 1000 to 1250 USD, this corresponds to 25% increase in price level (or inflation rate). If cost of a consumer basket increases from 1000 to 1100 USD, this corresponds to 10% inflation rate. An increase of cost from 1000 to 1050 USD corresponds to 5% inflation rate, and increase from 1000 to 1020 USD means 2% increase in price level.

Increase in the overall price level from April 2022 to April 2023:

- ◯ in percent ____

Display This Question:

If In your opinion, will prices in the U.S. be higher or lower in April 2023 relative to today? = ***Lower*** *than today*

C6a_2 How much do you think the **overall price level** in the U.S. will decrease between April **2022** and April **2023** (in percentage terms)?

For example, if cost of a typical consumer basket decreases from 1000 to 750 USD, this corresponds to 25% decrease in price level (or deflation rate, which is negative inflation rate). If cost of a consumer basket decreases from 1000 to 900 USD, this corresponds to 10% deflation rate. A decrease of cost from 1000 to 950 USD corresponds to 5% deflation rate, and decrease from 1000 to 989 USD means 2% decrease in price level.

Decrease in the overall price level from April 2022 to April 2023:

○ in percent ____

Page Break

Display This Question:

*If In your opinion, will prices in the U.S. be higher or lower in April 2023 relative to today? = About the **same** as today*

C6a_3 You have indicated that you expect that the overall price level in the U.S. will be about the same as today in April 2023. This could mean that the change equals zero percent or that the percent change is small. Please select a category that best describes your opinion.

- ○ In April 2023 by 5% lower than today
- ○ In April 2023 by 4% lower than today
- ○ In April 2023 by 3% lower than today
- ○ In April 2023 by 2% lower than today
- ○ In April 2023 by 1% lower than today
- ○ In April 2023 exactly the same as today
- ○ In April 2023 by 1% higher than today
- ○ In April 2023 by 2% higher than today
- ○ In April 2023 by 3% higher than today
- ○ In April 2023 by 4% higher than today
- ○ In April 2023 by 5% higher than today

End of Block: C. Prior - Forecasts

Start of Block: D. Task

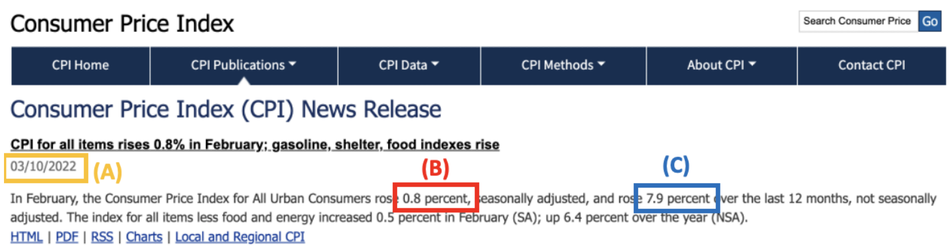
D

Recording Official Statistics

In the previous question, you answered that the overall price level in the U.S. will **change** by **${Your answer}%** over the next 12 months**.**

Next, we will ask you to fill a table with **official statistics** about the price level changes.

Based on the information from this screenshot, please fill the table below it.



Source: https://www.bls.gov/cpi/news.htm

Table

	Date of the news report	CPI inflation rate	
	mm/dd/yyyy (A)	in March 2022, in percent (B)	over the last 12 months, in percent (C)
Your answer			

Da You entered the following data based on the information from the screenshot:

Showing their transcription

If any data entry above is incorrect, please go back and enter correct information. Otherwise, proceed to the next questions.

We will NOT approve your HIT if you record the numbers from the screenshot *incorrectly*.

D2 According to the data you just entered, over the past 12 months, the overall price level in the U.S. has

- ○ decreased by 8.5%.
- ○ decreased by 1.2%.
- ○ not changed.
- ○ increased by 8.5%
- ○ increased by 1.2%

End of Block: D. Task

Start of Block: E. Posterior - Forecasts

E
Instructions:

Some of the following questions will ask you to forecast a change of a variable in the future in percentage terms (in other words, to provide your estimate of its growth rate).

For example, if the question asks about percentage change of average temperature in February 2023 relative to today and you think that it will be by 10% warmer in February 2023 than in February 2022 (i.e., the temperature will increase), enter “10.” If you think it will be by 10% colder in February 2023 than in February 2022 (i.e., the temperature will decrease), enter “-10”. If you think it will be about the same, enter “0.”

E1
After learning about the official statistics, by how much do you think the **overall price level** in the U.S. will change over the **next 12 months** relative to today (in percentage terms)?

If you think the overall price level will increase, enter a positive number. If you think it will decrease, then enter a negative number. If you think that the price level will not change, enter 0.

○ Price change over 12 months, percent ____

Display This Question:

If If After learning about the official statistics, by how much do you think the overall price level in the U.S. will change over the next 12 months relative to today (in percentage terms)? Response Is Equal to 0

E1_a You have indicated that you expect that the overall price level in the U.S. will be about the same as today in 12 months. This could mean that the change equals zero percent or that the percent change is small. Please select a category that best describes your opinion.

○ In April 2023 by 5% lower than today

○ In April 2023 by 4% lower than today

○ In April 2023 by 3% lower than today

○ In April 2023 by 2% lower than today

○ In April 2023 by 1% lower than today

○ In April 2023 exactly the same as today

○ In April 2023 by 1% higher than today

○ In April 2023 by 2% higher than today

○ In April 2023 by 3% higher than today

○ In April 2023 by 4% higher than today

○ In April 2023 by 5% higher than today

E2 By how much do you think the **average hourly earnings** in the U.S. will change **over the next 12 months** (in percentage terms)?

If you think the average hourly earnings will increase, enter a positive number. If you think they will decrease, then enter a negative number. If you think that the average hourly earnings will not change, enter 0.

○ Change in the average hourly earnings over the next 12 months, percent
____

Display This Question:

If If By how much do you think the average hourly earnings in the U.S. will change over the next 12 mon... Text Response Is Equal to 0

E2_a You have indicated that you expect that the average hourly earnings in the U.S. will be about the same as today in 12 months. This could mean that the change equals zero percent or that the percent change is small. Please select a category that best describes your opinion.

○ In April 2023 by 5% lower than today

○ In April 2023 by 4% lower than today

○ In April 2023 by 3% lower than today

○ In April 2023 by 2% lower than today

○ In April 2023 by 1% lower than today

○ In April 2023 exactly the same

○ In April 2023 by 1% higher than today

○ In April 2023 by 2% higher than today

○ In April 2023 by 3% higher than today

○ In April 2023 by 4% higher than today

○ In April 2023 by 5% higher than today

E3 What is your own forecast for the **Air Quality Index** in Seattle, USA **in 2 weeks**?

◯ Good (AQI 0-50)

◯ Moderate (AQI 51-100)

◯ Unhealthy for sensitive groups (AQI 101-150)

◯ Unhealthy (AQI 151-200)

◯ Very unhealthy (AQI 201-300)

◯ Hazardous (AQI 301-500)

E4 What share of the U.S. population will be **fully vaccinated** by the end of **May 2022**?

Fully vaccinated *means a person has received their primary series of COVID-19 vaccines (i.e. at least two doses of Moderna or Pfizer Biotech OR at least one dose of Johnson & Johnson's).*

____

E5 What do you think the **unemployment rate** in the U.S. will be in **April 2023** (in percent)?

Note: In February 2020, right before the pandemic, the unemployment rate was 3.5%. In April 2020 after the pandemic, the unemployment rate peaked at 14.7%.

◯ unemployment rate in April 2023 ____

End of Block: E. Posterior - Forecasts

Start of Block: F. Posterior Wage

F1 Suppose in the future we offered you to perform a *similar task* you did today (but without numerical literacy questions) taking about **10 min** of your time once a month. I.e., you would record the information from the same website and provide your prediction based on it.

How many **months** would be you interested in working?

- ◯ 0
- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

NOTE **WE MAY USE YOUR ANSWER TO THIS QUESTION TO OFFER YOU WORK ON FOLLOW-UP HITS.**

F2 In the previous question, you answered that you are willing to work on a *similar* 10-min task for ${your answer in F1} months, which corresponds to **$e{ 10 * your answer in F1}** minutes of your time. What is the **lowest** *total* reward that you would accept to work? (in USD)

- $e{ 0.4 * your answer in F1}
- $e{ 0.5 * your answer in F1}
- $e{ 0.55 * your answer in F1}
- $e{ 0.6 * your answer in F1}
- $e{ 0.65 * your answer in F1}
- $e{ 0.7 * your answer in F1}
- $e{0.75 * your answer in F1}
- $e{ 0.8 * your answer in F1}
- $e{ 0.85 * your answer in F1}
- $e{ 0.9 * your answer in F1}
- $e{ 1 * your answer in F1}
- $e{ 1.05 * your answer in F1}
- $e{ 1.1 * your answer in F1}
- $e{ 1.15 * your answer in F1}
- $e{ 1.2 * your answer in F1}
- $e{ 1.25 * your answer in F1}
- $e{ 1.3 * your answer in F1}
- $e{ 1.35 * your answer in F1}
- $e{ 1.45 * your answer in F1}
- $e{ 1.5 * your answer in F1}

- $e{ 1.6 * your answer in F1}
- $e{ 1.7 * your answer in F1}
- $e{ 1.8 * your answer in F1}
- $e{ 1.9 * your answer in F1}
- $e{ 2 * your answer in F1}
- Below $e{ 0.4 * your answer in F1}
- Above $e{ 2 * your answer in F1}

NOTE **WE MAY USE YOUR ANSWER TO THIS QUESTION TO OFFER YOU WORK ON FOLLOW-UP HITS.**

Display This Question:

*If In the previous question, you answered that you are willing to work on a similar 10-min task for... != Below $e{ 0.4 * your answer in F1}*

*And In the previous question, you answered that you are willing to work on a similar 10-min task for... != Above $e{ 2 * your answer in F1}*

F3 Would you be willing to accept an offer to do ${your answer in F1} ten-minute HITs that pay you total amount of $e{your answer in F2 + 0.05} USD?

◯ Yes

◯ No

Display This Question:

If Would you be willing to accept an offer to do ${your answer in F1} ten-minut... = No

*Or In the previous question, you answered that you are willing to work on a similar 10-min task for... = Below $e{ 0.4 * your answer in F1}*

*Or In the previous question, you answered that you are willing to work on a similar 10-min task for... = Above $e{ 2 * your answer in F1}*

F3_1 What is the smallest reward you would accept for ${your answer in F1} ten-minute HITs (total $e{ 10 * your answer in F1} minutes of your time)? (in USD)

◯ The smallest reward you would accept ________________________________

F4 What is the **smallest reward** for a **10-min** HIT you would **accept** for a *similar task* you did today in the next month?

- 0.00 - 0.50
- 0.51 - 0.60
- 0.61 - 0.70
- 0.71 - 0.80
- 0.81 - 0.90
- 0.91 - 1.00
- 1.01 - 1.10
- 1.11 - 1.20
- 1.21 - 1.30
- 1.31 - 1.40
- 1.41 - 1.50
- 1.51 - 1.60
- 1.61 - 1.70
- 1.71 - 1.80
- 1.81 - 1.90
- 1.91 - 2.00
- I would NOT accept any HIT that pays below 2.0 USD

Display This Question:

If What is the smallest reward for a 10-min HIT you would accept for a similar task you did today in... = I would NOT accept any HIT that pays below 2.0 USD

F5 What is the **smallest reward** you would **accept** for a 10-minute HIT *similar* to this one in the next month?

◯ Pay for 10 minutes, USD ____

Display This Question:

If What is the smallest reward for a 10-min HIT you would accept for a similar task you did today in... != I would NOT accept any HIT that pays below 2.0 USD

F6 You answered that you would accept **${your answer in F4}** **USD** per 10-min session for a *similar task* you did today in the next month. Please specify the smallest amount that you would accept to work.

◯ The smallest amount you would accept to work ____

End of Block: F. Posterior Wage

Start of Block: G. Qualification and experience-related questions

G This is the last group of short questions. It refers to you and your work experience.

G1 Think about the amount of time you devote to work on MTurk. Is this more or less than 20 hours per week?

- ○ More than 20 hours per week
- ○ Less than 20 hours per week

G1a How many hours do you work on MTurk in a typical week?

____

G2 Do you work on other crowdsourcing platforms in addition to MTurk?

- ○ Yes, regularly
- ○ Yes, occasionally
- ○ No

Display This Question:

If Do you work on other crowdsourcing platforms in addition to MTurk? != No

G2a How many hours per week do you usually work on other online platforms?

____

G3 Do you have a day job in addition to MTurk?

- ◯ Yes, a full-time job
- ◯ Yes, a part-time job
- ◯ No, but I am looking for one
- ◯ No, and I am not interested in getting another job

Display This Question:
If Do you have a day job in addition to MTurk? = Yes, a full-time job
Or Do you have a day job in addition to MTurk? = Yes, a part-time job

G3a How many hours per week do you usually work on day job(s)?

- ◯ <5
- ◯ 5-10
- ◯ 10-20
- ◯ 20-30
- ◯ 30-40
- ◯ 40 or more

Display This Question:
If Do you have a day job in addition to MTurk? = Yes, a full-time job
Or Do you have a day job in addition to MTurk? = Yes, a part-time job

G3b You have selected that you work ${your answer in G3a} hours a week. Please specify the average hours you usually work per week on day jobs.

- ◯ average hours you work per week ________________

Display This Question:

If Do you have a day job in addition to MTurk? = Yes, a full-time job

Or Do you have a day job in addition to MTurk? = Yes, a part-time job

G3c If you could choose the number of hours you work each week, and taking into account how that would affect your income, how much would you choose to work in **May 2022**?

- ◯ fewer hours than today
- ◯ about the same hours
- ◯ more hours than today

Display This Question:

If If you could choose the number of hours you work each week, and taking into account how that woul... = fewer hours than today

Or If you could choose the number of hours you work each week, and taking into account how that woul... = more hours than today

G3d How many hours a week would you choose to work on average in **May 2022**? Again, take into account how that would affect your income.

- ◯ Desired work hours in May 2022 ____

Display This Question:

If Do you have a day job in addition to MTurk? = Yes, a full-time job

Or Do you have a day job in addition to MTurk? = Yes, a part-time job

G3e1 What do you think is the percent chance that **four months from now** you will be...

Please enter a percent 0-100 for each. If you are certain that some event is impossible (e.g. you start your own business), answer 0.

Employed with the same employer : _______
Employed with a different employer : _______
Self-employed : _______
Unemployed and actively looking for a new job : _______
Not employed and not looking for a new job : _______
Total : ________

Display This Question:

If Do you have a day job in addition to MTurk? = Yes, a full-time job

Or Do you have a day job in addition to MTurk? = Yes, a part-time job

G3f1 Suppose someone **offered you a job** in **May 2022** in line with your current work that **pays by 10% more** than your current job. Would you accept this offer?

- Yes
- No
- Don't know

Display This Question:

If Do you have a day job in addition to MTurk? = Yes, a full-time job

Or Do you have a day job in addition to MTurk? = Yes, a part-time job

G3f11 What is the **smallest** increase relative to your current pay should a new job offer for you to **accept** it in May 2022?

- 0-2%
- 2-5%
- 5-7%
- 7-10%
- 10-15%
- 15-20%
- 20-25%
- 25-30%
- >30%
- I am not interested in another job

Display This Question:

If Do you have a day job in addition to MTurk? = No, but I am looking for one

G3e2 What do you think is the **percent chance** that **four months from now** you will be...
Please enter a percent 0-100 for each. If you are certain that some event is impossible (e.g. you start your own business), answer 0.

Employed : _______
Self-employed : _______
Unemployed and actively looking for a job : _______
Not employed and not looking for a job : _______
Total : ________

Display This Question:

If Do you have a day job in addition to MTurk? = No, but I am looking for one

G3f2 Suppose someone **offered you a job** in **May 2022** in line with your previous work. What the **smallest pay** should a new job offer **relative to your previous pay** for you to **accept** it?

- ○ by 15% or more lower than previous pay
- ○ 10-15% lower
- ○ 7-10% lower
- ○ 5-7% lower
- ○ 2-5% lower
- ○ 0-2% lower
- ○ same as previous pay
- ○ 0-2% higher
- ○ 2-5% higher
- ○ 5-7% higher
- ○ 7-10% higher
- ○ 10-15% higher
- ○ > 15% higher

Display This Question:

If Do you have a day job in addition to MTurk? != No, and I am not interested in getting another job

And Do you have a day job in addition to MTurk? != No, but I am looking for one

G5 In what industry is your main job?

- ◯ Agriculture, Forestry, Fishing or Hunting
- ◯ Mining, Quarrying, or Oil and Gas Extraction
- ◯ Utilities
- ◯ Construction
- ◯ Manufacturing
- ◯ Wholesale Trade
- ◯ Retail Trade
- ◯ Transportation or Warehousing
- ◯ Information Services (including Publishing or Media)
- ◯ Banking, Finance, or Insurance
- ◯ Real Estate, or Rental & Leasing Services
- ◯ Professional, Technical, or Business Services
- ◯ Education
- ◯ Health Care or Social Assistance
- ◯ Arts, Entertainment, or Recreation
- ◯ Hotel, Accommodation, Restaurant, or Food Services
- ◯ Other Services (except Government)
- ◯ Government, including Military
- ◯ Other: ____

Display This Question:

If Do you have a day job in addition to MTurk? = No, and I am not interested in getting another job

G5a Why are you not interested in getting a day job?

- ◯ I earn enough online (1)
- ◯ I need flexible schedule due to caregiving responsibilities (2)
- ◯ I am retired (3)
- ◯ I am a student (4)
- ◯ Due to health concerns or disability (5)
- ◯ Other: (6) ____

G6 What is your highest education level?

- ◯ Less than high school
- ◯ High school graduate
- ◯ Some college
- ◯ 2 year degree
- ◯ Bachelor's or other 4 year degree
- ◯ Master's or Professional degree
- ◯ Doctorate/PhD

G7 How often during the usual week do you check news?

- ○ I don't usually read/watch news
- ○ Every day
- ○ Almost every day
- ○ A few days

G8a What is your gender?

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)
- ○ Prefer not to say (4)

G8b How old are you?

____

G8c In which U.S. state do you currently reside?

(Multiple choice questions/ omitting options)

G8d What is your ethnicity?

- ☐ White
- ☐ Black or African American
- ☐ Hispanic or Latino
- ☐ Asian
- ☐ American Indian or Alaska Native
- ☐ Native Hawaiian or Pacific Islander
- ☐ Other
- ☐ Prefer not to answer

G9 Are you currently married or cohabiting?

- ○ Yes
- ○ No

Display This Question:

If Are you currently married or cohabiting? = No

G10 Have you ever been married?

- ○ Yes
- ○ No

G11 How many children under 18 do you have?

- ◯ None
- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5
- ◯ More than 5
- ◯ Prefer not to answer

G12 What is your annual income?

- Less than $10,000
- $10,000 - $19,999
- $20,000 - $29,999
- $30,000 - $39,999
- $40,000 - $49,999
- $50,000 - $59,999
- $60,000 - $69,999
- $70,000 - $79,999
- $80,000 - $89,999
- $90,000 - $99,999
- $100,000 - $149,999
- $150,000 - $199,999
- More than $200,000
- Prefer not to answer

G13 Can you recall how much have you spent on following products last month?

	Monthly Spending
	In USD
Food (including grocery, beverages, dining-out, take-out food, etc.)	
Gasoline	

G14 Was it confusing to answer any questions or to complete any tasks in this HIT? If so, please explain.

____

Completion

Your completion code is ${e://Field/compcode}.

End of Block: G. Qualification and experience-related questions