Macroeconomic Transmission of Financial Crises: A Story of Aggregate Demand

Nicolas Mäder†

November 4, 2019

Click here for the latest version.

Abstract

How do financial crises fuel protracted episodes of muted labor demand? The 2008 Financial Crisis caused a significant and long-lasting contraction in employment as firms persistently curbed their demand for labor. To rationalize such an extraordinary real response to financial crises, contemporary macroeconomic theory predominantly appeals to mechanisms of aggregate supply. In contrast, since anecdotal evidence suggests that the Great Recession was primarily fueled by a slump in consumer demand, I propose a macroeconomic framework in which financial crises cause real contractions via aggregate demand. In the model, financial crises arise when broker-issued margin calls force institutional investors — pension funds who allow households to indirectly hold claims in firms — to liquidate assets in a fire sale. Since fire sales occur at prices below fundamental values, they cause potentially substantial losses in nominal household wealth. To make up for lost retirement savings, households react to fire sales by reducing nominal consumer spending. In response to the sudden slump in aggregate demand, firms scale back employment and investment because worker effort is sensitive to nominal wage cuts. The resulting collapse of the capital stock causes a protracted depression of labor productivity and, in turn, a long-lasting labor market hysteresis. Since the real contraction’s severity and longevity increases in the size of the initial nominal collapse, I find that more aggressive unconventional monetary policy leads to less severe macroeconomic transmissions.

Keywords: Financial crisis, fire sale, aggregate demand, downward wage rigidity, involuntary unemployment, heterogeneous agents

JEL codes: E24, E44, E50, E70

†I am grateful to Peter Rousseau, Atsushi Inoue, Robert Driskill, Sheila Tschinkel, Frank Ciarliero, Daniel Manigrum, Mat Knudson, Tam Bui, Siraj Bawa, Jason Campbell, Sebastian Tello-Trillo, Kathleen McKiernan, Riyuichiro Izumi, You Du, Lingsheng Meng, Hesuk Jung, André Suter, Frank Bratschi and seminar participants at the I85 Macro Workshop, Western Economic Association International Annual Conference, Missouri Valley Economic Association Annual Conference, and Swiss Economists Abroad Annual Conference for insightful discussions and helpful comments. I would also like to thank Vanderbilt University for its generous support in the form of a Kirk Dornbush Summer Research Grant and the Walter Noel Dissertation Fellowship. Department of Economics, Vanderbilt University; nicolas.a.maeder@vanderbilt.edu
1 Introduction

How do financial crises fuel protracted episodes of muted labor demand? Over the course of the Great Recession, US employment fell by roughly 5% as seven million US workers were separated involuntarily and did not successfully find reemployment soon thereafter. The ensuing recovery period spanned roughly a decade, a phenomenon that has become known as labor market “hysteresis” (Yagan, 2019). The starting point of this paper is that the described, extraordinary labor market episode was caused by the preceding financial crisis. In particular, it has been established that the 2008 Financial Crisis significantly and negatively affected employment via both aggregate demand (Mian and Sufi, 2014) and aggregate supply (Chodorow-Reich, 2014), but contemporary macroeconomic theory predominantly emphasizes the latter. This paper’s main contribution thus lies in the formalization of a wealth effect driven aggregate demand channel in the context of macroeconomic general equilibrium.

My principal empirical objective is to match the observed, extensive labor margin since 2008. I focus on the extensive margin because it serves as a natural first-order approximation of the degree to which financial crises affect households heterogeneously. For example, consider Figure 1 which depicts the evolution of decomposed US unemployment since 1987.

Figure 1. Decomposed stock of unemployed US workers, 1987-2019

Notes: Figure 1 displays the longitudinal evolution of decomposed US unemployment since 1987. The fact that roughly seven million US workers who were separated involuntarily did not successfully find re-employment between 2008 and 2010 yields two main insights. First, the observed contraction in employment reflects a broad slump in labor demand. Second, the crisis did not affect all households uniformly as some workers involuntarily lost their job while others did not. As stock variables, the displayed series effectively represent integrals over all corresponding past inflows (e.g. quits, layoffs, labor force entry) and outflows (e.g. hires, labor force exit). Figure 1 thus reveals the Great Recession’s long-term effects that remain hidden under the often cited quits, layoffs, and discharges measure. All data was taken from the Bureau of Labor Statistics (BLS) with job leavers and job losers (laid off + others) relabeled as voluntarily and involuntarily unemployed.
The fact that between 2008 and 2010, roughly seven million US workers were separated invol-
untarily and did not successfully find re-employment yields two main insights. First, the decline in
employment observed during the Great Recession was driven by a broad slump in labor demand.\(^1\) Second, the crisis did not affect all households uniformly because some workers lost their employ-
ment involuntarily while others did not. Further notice that at any given time, only a small fraction
of unemployed US workers report to choose unemployment voluntarily, but that the contrast be-
tween voluntary and involuntary unemployment was particularly stark during the Great Recession.
Therefore, if financial crises cause extraordinary economic downturns and the corresponding rise
in unemployment is overwhelmingly involuntary, the welfare implications of financial crises are
particularly severe (see Shapiro and Stiglitz, 1985).

To formulate integrated policy recommendations for an economy that is subject to recurring
crises, I embed both crisis origination and crisis transmission into a unified macroeconomic frame-
work. Specifically, crisis origination is modeled explicitly because recognizing potential sources
of financial fragility is instrumental in the process of ex ante crisis prevention. In turn, the fire
sales are embedded in a macroeconomic model because we are not principally interested in the
nominal crises themselves, but rather in their real effects. Understanding the relevant transmission
mechanisms thus helps policy makers enact a sensible crisis response when a crisis is imminent.

Methodologically, I generate financial crises by way of a time-invariant, but occasionally binding
maintenance margin which causes, if binding, a nominal collapse of the economy’s numéraire.\(^2\) In
terms of the ensuing transmission, I break with the traditional approach of appealing to aggregate
supply and instead propose the following wealth-effect-driven demand channel. First, during a
financial crisis, households lose a potentially substantial fraction of their accumulated retirement
savings. Second, to make up for the incurred losses in projected retirement wealth, households
proceed by reducing consumer spending. Third, firms react to the upcoming slump in consumer

\(^1\) Notice that a falling labor demand does not require an increase in corporate layoffs. In fact, “notwithstanding a
spike in 2008”, separation only accounted for roughly one quarter of the observed variation in unemployment during
the Great Recession (Shimer, 2012). To understand this, consider the following thought experiment. Suppose there
is an exogenous, time-invariant separation rate. In this scenario, firms have to hire a certain number of workers each
period just to retain a constant work force. To shrink their labor force during a downturn — when labor demand falls — firms then only need to reduce the number of new hires without ever having to lay off any workers.

\(^2\) A maintenance margin is the minimum amount of equity required to maintain a margin account with a brokerage
firm. If a declining market value of an investor’s portfolio causes equity to fall below the maintenance threshold,
the broker issues a margin call. A margin call is a broker-issued liquidity demand with the aim of consolidating the
corresponding investor’s equity position.
demand by scaling back production via employment (and investment).³

To incorporate occasional financial crises into a macroeconomic framework, I propose the following key model elements. First, consumption goods producers must finance labor and capital rentals with commercial loans. Since production occurs prior to the realization of consumer demand, commercial loans are subject to credit risk. Banks follow an originate-to-distribute model with commercial loans being sold, securitized, and ultimately held by pension funds in the form of a collateralized loan obligation (CLO). In turn, an occasionally observed, fundamental demand signal yields a new secondary market CLO price with pension fund assets being marked-to-market accordingly. If a pension fund’s resulting equity position violates the prevailing maintenance margin, the broker issues a margin call, which induces, if liquidity is low, a fire sale.⁴ Following the fire sale, households make up for lost retirement savings by curbing consumer spending.

To generate real effects from the nominal shock, I exploit the well established finding that wages are nominally downward rigid.⁵ Specifically, I follow Solow 1979 and Akerlof and Yellen’s (1990) in assuming that exerted worker effort is sensitive to nominal wage cuts, in which case nominal demand slumps cause declines in demand for labor and investment. Recall, however, that beyond a sudden and sharp increase in unemployment, the Great Recession also featured a long and protracted labor market “hysteresis” (Yagan, 2019). To match the observed recovery period, I require an additional friction. For this, the labor market setup is borrowed from Weiss (1980) where information asymmetries between firms and workers give rise to a real downward wage rigidity which generates involuntary unemployment that persists. Specifically, since the attractiveness of each worker’s outside option is increasing in their skill, lowering wage offers always induces the highest skilled employees to quit first. Depending on the cross-sectional distribution of skill, adverse selection may then induce a threshold below which effective labor costs are decreasing in the wage. As a result, firms would rather face excess labor supply than paying wages below the threshold (Weiss, 1980).

---

³Since production and sales are separated intratemporally in my model, firms rent capital and hire labor while demand is still uncertain.

⁴Building on the extensive intuitive account provided by Shleifer and Vishny (2011), a fire sale is understood as a forced placement of sell orders irrespective of the corresponding asset’s fundamental value. Institutionally, fire sales thus constitute an example of a “portfolio-adjusting” trade, whereas orders based on new fundamental information may be described as “information-motivated” (Cuneo and Wagner, 1975). Margin calls serve as a natural way to induce investors to sell at a price at which they would not normally — in absence of the margin call — want to sell.

⁵See Fallick, Lettau, and Wascher (2016), Daly, Hobijn, and Lucking (2012), and Daly, Hobijn, and Wiles (2011) for recent empirical evidence that wages are nominally downward rigid.
To pin down the empirically relevant recovery duration to be matched by the model, I follow Yagan (2019) in plotting a set of labor market measures relative to their respective pre-crisis levels at the beginning of 2008.

**Figure 2.** Labor market recovery following the financial crisis of 2008

![Labor market recovery following the financial crisis of 2008](image)

Notes: Figure 2 depicts several different labor market measures following the shock in 2008. While unemployment had reverted back to pre-crisis levels by late 2015, employment recovery has been much more sluggish as many workers exited the labor force during the Great Recession. On the other hand, the displayed discrepancy is also due to demographic change as illustrated by the employment series that conditions on working age. The latter will then serve as the relevant labor market measure to be matched because demographic trends and labor force participation are held constant in my theory. All series were taken from BLS and normalized to 2008 levels.

Notice that while the conventional metric of unemployment had reverted back to pre-crisis levels by late 2015, recovery in employment occurred much more slowly. In fact, as of 2019, the employment-to-population ratio has still only recovered half-way since its most recent trough in 2010. Both of the aforementioned metrics, however, are susceptible to changes in the labor force. In fact, focusing on the restricted subsample of workers between 25 and 54, we can see that virtually all of the remaining employment-to-population gap is attributable to aging. Since I abstract from demographic change in the model, the empirically relevant target for labor market recovery is roughly ten years.\(^6\)

The methodological contribution of the paper is twofold. First, the proposed framework rationalizes the statistics recorded during the Great Recession thereby elevating the corresponding ergodic tail out of the realm of statistical ‘outliers’. To illustrate this, consider Figure 3 which plots US unemployment against time and against its state space in the form of a set of estimated ergodic densities.

\(^6\)Holding the labor force constant implies that any change in unemployment results in an equivalent change in employment (and vice versa). Since measured unemployment is not invariant to changes in the labor force, the targeted recovery duration must account for population aging as does the working age employment recovery series depicted in Figure 2.
Figure 3. Decomposition of unemployment in crisis and non-crisis episodes

Notes: Figure 3 depicts three (normalized) estimated ergodic densities for US unemployment in the post-Volcker era. The unconditional density illustrates that unemployment exhibits a substantial upper tail with the two conditional densities confirming that said tail was recorded following the 2008 Financial Crisis. For purposes of partitioning the data, the “labor market crisis” episode is constructed by collecting all quarters that featured an unemployment rate higher than the previous peak in 1992. All density estimates were derived using the Gaussian kernel proposed in Botev et al. (2010) with mesh size $2^{-7}$. The data was recovered from BLS.

From the standpoint of economic intuition, it is unsurprising that unemployment reached its peak during the Great Recession. But if this is the case, we may conjecture that, as a general principle, if our economy of interest is subject to occasional crisis episodes, the data will be skewed, fat tailed, or, if downturns are particularly severe, even multimodal.\(^7\) In the spirit of Stiglitz (2018)\(^8\), one methodological contribution of this paper then lies in the joint rationalization of both the statistics recorded during crisis episodes as well as the statistics recorded during non-crisis episodes.

As a second methodological contribution, the proposed framework illustrates the high degrees of state and parametric heterogeneity that can be accommodated when model primitives are chosen subject to the cognitive constraint that agents are incapable of solving Euler equations in an internally consistent manner. In particular, I argue that even if our theory produces decision rules that appear appealing intuitively and match the data, but no real-world agent is realistically able to derive them, the modeler for all intents and purposes imposes behavior as the model primitive and thus invariably obfuscates the actual tradeoffs considered by those agents. In effect, internal

\(^7\)While Kocherlakota (2000) emphasizes the importance of exploring the business cycle’s cyclic asymmetry in general terms, Brunnermeier and Sannikov (2014) specifically argue that occasional crises episodes may give rise to ergodic multimodality. In either case, rationalizing the observed dynamics requires a strategy, theoretical and/or empirical, that extends beyond matching first and second data moments.

\(^8\)“The most important challenge facing any macro-model is to provide insights into the deep downturns that have occurred repeatedly.”
consistency only serves as a useful benchmark if we can realistically assert that agents’ behavior indeed derives from the proposed optimization problem. For example, a household considering the marginal cost of increasing consumption today realistically resorts to quantifying those costs in terms of lost (retirement) savings, not in terms of future consumption. This is because (i) an accurate probabilistic assessment of future consumption is prohibitively expensive, (ii) accumulated savings serve as a store of value in the sense that future consumption is strictly increasing in accumulated savings, and (iii) it is quantitatively convenient.

The remainder of the paper is organized as follows: Section 2 places the paper in the literature by examining the proposed framework’s key elements. Section 3 presents the theoretical framework that formalizes the desired demand channel. Section 4 discusses the pursued parameterization strategy, illustrates quantitatively the macroeconomic transmission of a typical fire sale episode, and assesses various monetary policy options. Section 5 presents a summary of the gained topical and methodological insights. Section 6 concludes.

2 Incorporating fire sales into a macroeconomic framework

The present work’s objective most closely resembles the recent article of Gertler, Kiyotaki, and Prestipino (2017) (GKP), but the pursued approaches differ substantially. First, the financial crises embedded in the macro model considered here manifest themselves in the form of a fire sale, whereas in GKP they take the form of a bank run. Second, macroeconomic transmission occurs via a slump in consumer demand in my model, whereas in GKP bank runs temporarily prevent banks from financing investment with less efficient intermediaries (households) taking their place. Methodologically, the proposed crises are similar in the sense that both types arise from an occasionally binding constraint, but equilibrium here is unique, whereas it is indeterminate in GKP. I now turn to placing the paper in the literature more broadly by examining more closely the key elements — origination, transmission, and mitigation — of the simulated crises.

Appendices A through E contain a derivation of individual equilibrium strategies, a description of the employed data, a motivating discussion of the proposed consumption-savings problem as well as a technical note regarding the often cited notion of ergodicity, and an institutional dictionary.

While focusing on demand now, I ultimately envision a framework that unifies both demand and supply channels. The key point here is that while indeterminacy (and/or multistability) are conducive to generating extraordinary economic behavior, they are not necessary. Instead, my key methodological crisis component is that equilibrium is fragile in the sense that it is discontinuous (in the state space).
Crisis origination

How do the simulated financial crises originate? The origination mechanism considered here resembles the canonical fire sale in Brunnermeier and Pedersen (2009). Since the CLO is purchased on margin, binding maintenance margin requirements induce demand to be non-monotonic (see Figure 4). Once the price of the CLO falls below the critical threshold, the Walrasian method of countering excess supply by lowering the price effectively increases excess supply because brokers are forcing investors to liquidate larger parts of their portfolio. Maintenance margins thus gives rise to a price threshold below which there exists a “diabolic feedback loop” between falling asset prices and increasing margin calls (Brunnermeier and Pedersen, 2009). As such, the economic intuition underlying the proposed fire sale here is virtually equivalent to the one in Brunnermeier and Pedersen (2009) with the primary difference being that equilibrium here is unique because maintenance margins are fixed. In contrast, many articles in the literature model financial crises to manifest themselves in the form of equilibrium multiplicity.

Figure 4. The secondary market for CLOs

Notes: Figure 4 illustrates the effects of a prevailing maintenance margin. Since CLO demand is non-monotonic, an adverse price shock can spark a cascade of forced sell orders which causes excess supply to increase as the price falls.

---

12In such a scenario, if liquidity shortage is symmetric, private market equilibrium ceases to exist such that financial markets collapse altogether unless a third party intervenes. I thus assume that, once a fire sale is imminent, the central bank supplies additional liquidity by deploying unconventional monetary policy.

13While unique, equilibrium here is fragile in the sense that it is discontinuous in the state space. Since occasionally binding constraints typically induce such discontinuities, they require special attention when solving for equilibrium, but have been a subject of heightened interest since the 2008 financial crisis (see Cuba-Borda et al., 2019).

14Whether maintenance margins are set exogenously or endogenously is not paramount in terms of examining the macroeconomic transmission following a fire sale. Therefore, since I only require that fire sales in fact occur occasionally, the exogenous setting is preferred.

15Prominent examples other than Brunnermeier and Pedersen (2009) include the canonical bank run by Diamond and Dybvig (1983), the sovereign defaults by Calvo (1988). Lorenzoni and Werning (2019) go even further in that they examine an economy featuring both multiple equilibria and multiple steady states. In Gertler, Kiyotaki, and Prestipino (2017), equilibrium is unique but discontinuous in the state space.
**Crisis transmission**

Following origination, how are nominal crises transmitted to the real sector? If the premise of a causal link between nominal collapses and real downturns is accepted, one may wonder whether transmission chiefly occurs via aggregate supply or via aggregate demand. The supply view suggests that financial crises cause contractions because financing production becomes more expensive for firms. Conversely, the demand view holds that recessions rather reflect firms reacting to shifts in aggregate demand as households reduce their consumption spending. While both views imply a contraction in aggregate production, the resulting effects on prices are opposite. Since the large declines in output observed during the Great Recession were not accompanied by a corresponding price shift in either direction, anecdotal evidence suggests that supply and demand channels coexist, an assertion that has been affirmed empirically. Focusing on housing net worth, Mian, Rao, and Sufi (2013) find that the recent financial crisis caused a severe slump in consumer demand. Weak aggregate demand in turn had a significant negative impact on employment (Mian and Sufi, 2014). Conversely, Chodorow-Reich (2014) successfully establishes the existence of a supply channel by exploiting bank-firm level data and estimates that at least one third of all employment losses at small to medium sized firms (SME) during the Great Recession are attributable to credit withdrawals by banks. While it is thus likely that both channels were operative, Figure 5 provides some additional anecdotal evidence that the relative contribution of aggregate demand was stronger.

**Figure 5.** “What is the single most important problem facing your business today?”

Notes: Figure 5 reports business owners’ answer to the question “What is the single most important problem facing your business today?”. The fact that business owners were more worried about poor sales than securing finance during the Great Recession suggests that the relative contribution of aggregate demand in creating the recession was higher than the contribution of aggregate supply mechanisms. The data was retrieved from the online appendix in Mian and Sufi (2014).
While Figure 5 does not imply that obtaining finance was easily accessible during the Great Recession, it does speak to the relative importance of aggregate supply and aggregate demand mechanisms because it illustrates that the primary concern of business owners during the Great Recession was that they would be unable to sell products and/or services, not that they would be unable to obtain financing.

Following the 20% decline in household wealth during the 2008 Financial Crisis, the demand-side literature has refocused on rationalizing the relatively stark, observed consumer responses to changes in wealth and income (Mian, Rao, and Sufi, 2013; Kaplan and Violante, 2014). For this, first notice that a fixed aggregate wealth loss that is concentrated among households with higher marginal propensities to consume (MPC) causes a larger slump in aggregate demand. The proposition that households may have heterogeneous MPCs goes back to Keynes (1936) who conjectured that the latter is decreasing in wealth. Zeldes (1989) and Carroll and Kimball (1996) show that the proposed concavity can be rationalized by pairing income uncertainty with a precautionary savings motive on part of the household. In turn, King (1994) emphasized that concavity can also arise if consumers face a binding borrowing constraint, an assertion reiterated by Carroll (2001): “for many purposes the behavior of constrained consumers is virtually indistinguishable from the behavior of unconstrained consumers with a precautionary motive”.

If the consumption function is concave and all wealth is stored in liquid form, all high MPCs are concentrated among households with lower wealth. In such a setting, it is difficult to explain the large consumer responses to temporary fiscal stimulus payments that have been observed in the past. Kaplan and Violante (2014) resolve this issue by arguing that even wealthy households can have high MPCs, namely when most of their wealth is tied up in illiquid investments. In their setup, wealthy households may voluntarily choose a “hand-to-mouth” strategy because purchasing and liquidating illiquid assets is costly. While Kaplan and Violante (2014) are interested in tracking the effects of loosening the household’s borrowing constraint via fiscal stimulus payments, my paper is closer in spirit to Guerrieri and Lorenzoni (2017) who conduct a quantitative exercise documenting the effects of a shift from “easy credit to tight credit”. Since borrowing capacities tighten exogenously in their model, the corresponding shock may be interpreted as a demand-side analogue to the “financial shock” experienced by firms in Jermann and Quadrini (2012).
The present approach departs from the aforementioned demand-side literature in two principal ways. First and foremost, my aggregate demand slump is generated via an endogenous, widespread loss in nominal household wealth, not via an exogenously tightening credit constraint or even the concavity of the policy function. In particular, because consumption is increasing in wealth and wealth collapses during financial crises, the consumption function’s second derivative is not of primary importance as long as it is monotonically increasing. Moreover, since my economy’s numéraire is given by money, fire sales only cause nominal wealth losses, which are by themselves insufficient to generate significant real downturns because prices and wages are principally capable of absorbing nominal shocks of an arbitrary size. I thus follow the New Keynesian tradition in assuming that wages are nominally downward rigid. Second, I follow the macroeconomic tradition of modeling firms explicitly, which allows for the desired investigation of the extensive labor margin.

In terms of macroeconomic theory, early accounts of the nominal-real nexus had focused on aggregate demand (Keynes, 1936) and the accompanying phenomenon of deflation (Fisher, 1933), but the contemporary literature has effectively been dominated by supply-side shocks — most notably technology shocks — since Kydland and Prescott (1977, 1982). For example, the canonical “financial accelerator” literature, adverse technology shocks are exacerbated by deterioration of firm net worth which disincentivizes production via unfavorable external financing conditions (Bernanke, 1983; Bernanke and Gertler, 1989; Bernanke, Gertler, and Gilchrist, 1999). Conversely, Jermann and Quadrini (2012) introduce perturbations that originate in the financial sector, namely shocks that alter a firm’s borrowing capacity holding fixed the latter’s existing stock of collateral. In their model, the financial sector thus not only acts as an amplifier, but it constitutes a source of macroeconomic fluctuations in its own right. While output contractions are caused by binding quantity constraints in Jermann and Quadrini (2012), Gertler, Kiyotaki, and Prestipino (2017) appeal to lending and investment frictions that manifest themselves in the form of increasingly expensive capital loans following a bank run. Specifically, financing production is more expensive during a crisis because households — who temporarily own the entire capital stock — are less efficient “in handling investments”.

\[^{16}\text{In fact, my households are unable to borrow altogether.}\]

\[^{17}\text{Since households incur a utility cost from holding capital beyond a certain threshold, they must be compensated to do so following a bank run. “The fire sale of assets from banks to inefficient households will lead to a sharp rise in the cost of credit, leading to an extreme contraction in investment and output.” Therefore, bank runs cause extreme recessions because financing production becomes more expensive for firms.}\]
In my framework, the key assumption that links financial crises to real downturns is that worker effort is sensitive to nominal wage cuts such that real wages inflate immediately and substantially following a fire sale. Figure 6 confirms that such an immediate and substantial real wage inflation took place in the United States during the Great Recession. Notice, however, that the protracted nature of the Great Recession is hard to reconcile with a rigidity that is purely nominal (Elbsy et al., 2016). And indeed, in my model, the nominal rigidity only serves as a transmission mechanism initially. In turn, the empirically observed hysteresis is generated via a collapsing capital stock which induces a protracted decline in labor productivity.

**Figure 6.** Average hourly real wage and working-age employment in the US since 1987

![Chart showing the evolution of employment-to-working-age-population ratio and CPI-adjusted, average hourly wages in the post-Volcker era.](chart)

Notes: Figure 6 displays the evolution of the employment-to-working-age-population ratio and CPI-adjusted, average hourly wages in the post-Volcker era. During the Great Recession, real wages experienced an immediate and substantial spike while employment underwent a rapid and substantial decline. The proposed theory rationalizes this development as follows. Since worker effort is sensitive to nominal wage cuts, firms reacted to the slump in aggregate demand by curbing their demand for labor while leaving nominal wages largely unchanged. All data was retrieved from FRED.

The proposition that wages are nominally rigid is hardly new. Wage stickiness constitutes an integral part of the New Keynesian literature, but asymmetric downward rigidity has gained some additional momentum following the Great Recession. A classic early reference in this regard is the *General Theory*, in which Keynes challenges the prevailing view that unemployment was voluntary during the Great Depression because wages were nominally rigid (Tobin, 1972). Unsurprisingly, in light of the theory proposed herein, the Great Depression was later found to have been amplified if not caused by nominally downward rigid wages (Akerlof et al., 1996; Bernanke and Carey, 1996).

---

18 Classic examples of New Keynesian DSGE models with symmetric wage rigidities include Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). Asymmetric downward rigidities are examined by Kim and Ruge-Murcia (2011) and, more recently, by Schmitt-Grohé and Uribe (2016) and Na et al. (2018).

19 In this sense, moderate rates of inflation may serve as ‘grease’ for the labor market because it loosens the nominal wage constraints faced by firms (Kahneman et al., 1986).
More recently, Bakker (2015) replicates Bernanke and Carey’s (1996) cross-country exercise with Great Recession data and finds that rigid wages constitute a key driver of unemployment still.

From an empirical perspective, there is ample evidence that wages are downward rigid.\textsuperscript{20} In fact, since “the existence of wage stickiness is not in doubt” (Kahneman et al., 1986), it is unsurprising that there exists a subliterature that has interviewed firms — as wage setters — to elicit exactly why wages are downward rigid. In a survey of 184 firms, Campbell and Kamlani (1997) find that the two mechanisms exploited herein — endogenous worker effort and adverse selection — are in fact the two principal reasons why firms are hesitant to cut wages. Bewley (1999) and Blinder and Choi (1990) also find strong evidence in support of the worker effort hypothesis, but the latter find no evidence that firms fear adverse selection in hiring. They hypothesize, however, that adverse selection may play a larger role in quits, a conjecture confirmed by Campbell and Kamlani (1997). On the worker side, Kahneman et al. (1986) find that nominal wage cuts are often perceived as unfair, although there may be mitigating circumstances in case the employer faces bankruptcy.

In line with the firm-side results discussed above, theory has traditionally attributed downward rigidity to the prevalence of ‘efficiency wages’, which formalize the notion that labor productivity is increasing in the wage: “you get what you pay for” (Solow, 1979). To generate the desired dependence, Weiss (1980) pairs heterogenous labor productivity with asymmetric information to induce adverse selection, whereas Shapiro and Stiglitz (1984) give workers the ability to shirk. Potential other reasons for the prevalence of efficiency wages are costly labor turnover (Stiglitz, 1974) and endogenous effort as captured by Akerlof’s positive “gift exchange” mechanism (1982) and its negative “fairness” counterpart (Akerlof and Yellen, 1990).

\textit{Crisis response}

This paper also relates to the literature that examines how policy makers should optimally address an ongoing crisis. For this, I proceed by first discussing a select set of works that address the Great Depression, before examining more recent insights gained during the Great Recession.

The classical reference for the policy response during the Great Depression is \textit{A Monetary History of the United States} by Friedman and Schwartz (1963). The main thesis of their predominantly empirical exercise is to show that exogenous changes in monetary aggregates have real effects. In

\textsuperscript{20}See McLaughlin (1994), Akerlof et al. (1996), Card and Hyslop (1997), Lebow et al. (2003), Fehr and Goette (2005), Dickens et al. (2007), Daly et al. (2011), Daly et al. (2012), Barattieri et al. (2014), Fallick et al. (2016).
particular, they hypothesize that the Great Depression would have been less pronounced had the Federal Reserve acted less hawkishly.\footnote{Recall that the Fed iteratively raised rates between 1928 and 1932.} Romer and Romer (2010) echo the formers’ hypothesis, but find the proposed link to be lacking in its theoretical foundation. They thus proceed by formulating a concrete transmission story, namely that money affects the real sector via increasing real interest rates as prices fall. Therefore, assuming the premise that the real downturn during the Great Depression was caused by deflation is accepted, Romer and Romer (2010) establish the desired causal link by empirically showing that monetary policy indeed fueled deflation at the time. In line with the above references, Christiano, Motto, and Rostagno (2003) construct a quantitative DSGE framework and find — by generating counterfactuals for a set of differing monetary policy rules — that the Federal Reserve could have muted the Great Depression had it acted more dovishly.

Before delving into the literature on the Great Recession, notice that the aforementioned analyses all consider conventional interest rate policy as the central bank’s primary if not only monetary policy lever. But, as we observed during the Great Recession, solely relying on conventional measures may be severely limiting insofar as nominal rates are subject to the zero lower bound. However, as the ultimate creator of the numéraire, the central bank can intervene in many ways that go beyond traditional interest rate policy if necessary. Examples of unconventional tools include emergency lending facilities as well as outright purchases and repurchase agreements of non-governmental assets.\footnote{See Reis (Fall 2009) for a contemporaneous assessment of policy at the height of the Recession in late 2009.} In light of the widespread, recent deployment of such tools, it is unsurprising that researchers were quick to empirically examine and incorporate them in their models. Gertler and Karadi (2011) consider a DSGE framework, in which the central bank directly lends to the private sector when financial intermediaries are undercapitalized. They find that during crisis episodes, unconventional policy of this sort is beneficial irregardless of whether the zero lower bound binds, but that the benefits are particularly stark if it does bind. Similarly, Wu and Xia (2016) quantitatively assess the effects of unconventional policy on unemployment and find significant effects in the desired direction. In a more qualitative exercise, Kuttner (2018) concludes that the Federal Reserve’s actions were appropriate in the sense that the incurred costs are “dwarfed by the costs of the more protracted recession in the United States that likely would have occurred in the absence of the unconventional policies”.
3 A recursive sequential game

I start by providing an intuitive overview of the proposed institutional setup. First, households live for a finite number of periods. As members of the working age population, households supply labor to the corporate sector, whereas, following retirement, they proceed by drawing down their retirement savings. The latter are accumulated as workers entrust part of their labor income to financial intermediaries each period. Since each period evolves sequentially, the same household faces multiple optimization problems per period. Lastly, households are heterogeneous in both their states and in their parameterization.\textsuperscript{23}

The private sector consists of consumption goods producers, capital goods producers, and financial intermediaries. Capital firms produce, own, and rent out capital to the consumption goods sector. In turn, consumption goods firms use capital and labor to produce and sell consumption goods to households. Since sectoral competition occurs via quantity in the consumption goods sector, sectoral size serves as a natural lever to generate default via a calibration of expected profits. Banks scale up and down the economy’s artificial numéraire by issuing and subsequently selling commercial loans. Nonbank financial institutions (NBFI: pension funds and brokerage firms) allow households to invest their nominal wealth in the form of equity and debt contracts respectively.

The public sector consists of a central bank and a government. The central bank enacts conventional monetary policy by periodically adjusting its interest rate target, which is perceived as ‘the current interest rate level’ by all agents. In addition, if a financial crisis is imminent, the central bank can, as the ultimate creator of the numéraire, deploy unconventional policy that extends beyond the periodic announcement and defense of the interest rate target. The government taxes labor income, disburses unemployment and retirement benefits, and issues a government bond.

The framework is sequential for two principal reasons. First, to generate commercial loan defaults, I require that consumer demand is uncertain when the factors of production — capital and labor — are contracted. Second, to generate fire sales, the supply of the CLO must be fixed.

\textsuperscript{23}Heterogeneity poses two main methodological challenges. First, when the state space is infinite dimensional, contemporary practice requires some sort of approximation in determining equilibrium. Of course, this problem is only exacerbated when, in addition to states, parameters vary by households as well. Second, identifying a parameter’s cross-sectional distribution requires micro level data. In this paper, the first issue is addressed by requiring that each agent’s optimization problem be trivial to solve numerically. Rather than solving for a policy function, I instead proceed by recursively resolving each agent’s optimization problem each period for the current state only. The second issue is beyond the scope of this paper.
Consider a dynamic closed economy populated by a set of infinitely many households $J^{HH}$ of time-invariant measure $\mu^{HH} \equiv \mu(J^{HH})$. Time is discrete with $t \in \mathbb{N}$ denoting a period. Each household born in period $t$ is member of generation $t$ and lives for a finite, predetermined number of periods $T^L$. At age $T^R < T^L$, households lose their productive labor endowment, exit the labor force, and finance their consumption via a government pension and accumulated retirement savings. After exiting the economy at age $T^L$, an old household’s remaining property is, if applicable, bequeathed to a new household who takes its place. All initial generations are of equal size such that the labor force $J^{LF} \subset J^{HH}$ is of time-invariant measure $\mu^{LF} = \mu^{HH}(T^L - T^R)$. The corporate landscape is made up of a consumption and a capital goods sector, each consisting of a time-invariant set of firms $J^C$ and $J^K$ respectively, as well as a set of banks $J^B$, pension funds $J^PF$, and brokerage firms $J^BR$. Each period unfolds sequentially with $S^t_j$ denoting the system’s state in subperiod $j \in J^T = \{0, 1, 2, 3, 4, 5\}$ of period $t$. Households are heterogenous such that $S^t_j$ is infinite dimensional. State evolution is governed by equilibrium mappings from $S^t_j$ to a vector of controls $X^t_j$ and subject to exogenous innovations $\varepsilon^t_j$. The economy’s stochastic environment, as captured by the joint density of $\{\varepsilon^t_j\}_{j \in J^T, t \in \mathbb{N}}$, is given by a probability space $(\Omega, \mathcal{F}, \mu)$. Figure 7 illustrates the described notion of intratemporal sequentiality.

Figure 7. State transition in a sequential setup

Notes: The above graphic illustrates the sequential nature of the recursive game. The endogenous state $S^t_j$ and the exogenous innovation $\varepsilon^t_j$ give rise to equilibrium as captured by the vector of controls $X^t_j$, which in turn yields the new endogenous state. Period $t$ period ends after subperiod $t_5$ at which point the new period $t + 1_0$ begins.

24 $\mu^{HH}$ acts as a scaler of the economy and is not assumed to be unity. If $x_{ht}$ denotes a household specific quantity, the corresponding economy-wide aggregate is given by $x_t = \int x_{ht} \, dh$. Equivalently, $x_t = \sum_{j \in J^C} x_{ft}$ and $x_t = \sum_{j \in J^K} x_{ft}$ denote sectoral aggregates in the corporate sector.

25 I assume that all of the working age population is in the labor force.

26 I further proceed by highlighting random variables that are, at the contextually relevant point in time, non-predetermined by $\hat{X}$ with $X^t_i$ denoting a specific agent’s projection thereof. For example, knowing aggregate sectoral output is insufficient to infer the equilibrium consumption goods price at $t_1$ because, at the time, the realization of the households’ taste shock has not been observed. $\hat{P}_t$ thus denotes the corresponding random variable implied by the model with the accent simply highlighting the fact that its realization is unknown when firms are devising their optimal strategies. Moreover, since firms resort to approximating concurrent consumer demand with an isoelastic function, $\tilde{P}^i_t$ denotes firm $i$’s price projection.
The six subperiods unfold as follows. At $t_0$, households partition their retirement portfolio into equity and debt while pension funds decide on their capital structure. During $t_1$, capital goods and consumption goods firms produce output while households supply labor across the two sectors. Consumption goods producers finance their production with a commercial loan originated and distributed by banks. At $t_2$, all commercial loans are pooled and securitized into a CLO that is held by pension funds. At $t_3$, the occasional observation of a noisy taste shock signal leads to an information-motivated repricing of the CLO. Pension fund equity is repriced as a residual of assets net of debt. If equity falls below the prevailing maintenance margin, a margin call is issued and debt must be repurchased. If pension fund liquidity is insufficient to satisfy the margin call, pension funds are forced to liquidate part of their portfolio in a fire sale. At $t_4$, the taste shock realization determines aggregate demand with market clearing determining the corresponding equilibrium price. Finally, during $t_5$, all claims are settled, the central bank announces a new risk free interest rate target, and the government auctions off a new issue of its bond. Figure 8 summarizes which markets are open in which subperiod.

**Figure 8.** Intratemporal timeline

Notes: Figure 8 displays each period’s six subperiods including the respective markets that are open at each time.

Before describing in detail the proposed intratemporal sequence of events, I discuss the relevant market and equilibrium concepts. Markets are highly incomplete with trading exclusively taking place in recurrent spot markets, most of which only allow access to certain types of agents. For example, households cannot purchase commercial loans or securities directly, but they may hold an indirect claim on such assets via shares of NBFIs. The relevant equilibrium concept is the following.
**Recursive general equilibrium.** A recursive model economy is said to be of the general equilibrium type (RGE) if the price in each market is determined endogenously by pseudo-Nash equilibrium in (each subperiod of) each period.\(^{27}\)

Whenever prices are not set by market participants themselves, I entrust price discovery to a Walrasian auctioneer with the objective \(\max_{p \in \mathbb{R}^+} zp\), where \(p\) is the price chosen by the auctioneer and \(z(p)\) denotes the corresponding market’s excess demand in optimum. I further follow Arrow and Debreu (1954) in assuming that the auctioneer does not recognize the functional relationship \(z(p)\). In effect, this implies that excess supply induces the auctioneer to lower the price, whereas excess demand induces the auctioneer to raise the price. General equilibrium thus requires \(z(p^*) = 0\) in all Walrasian markets.\(^{28}\)

Table 1 summarizes all of my economy’s intratemporal markets. In each market not labeled as Walrasian, pricing is determined by the designated agents. Notice that even in absence of a Walrasian auctioneer, competition may still lead to market clearing as exemplified by the markets for Fed Funds, commercial loans, capital, and the consumption good. Contrarily, in spite of competition, the market for labor may not clear, even in equilibrium.

**TABLE 1. MARKETS AND PRICES**

<table>
<thead>
<tr>
<th>Price</th>
<th>Pricing</th>
<th>Clearing</th>
<th>Relative</th>
<th>Realized</th>
<th>Determined</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R^{FFR}_t) Fed Funds Rate (FFR)</td>
<td>Banks</td>
<td>Yes</td>
<td>Yes</td>
<td>(t_1)</td>
<td>(t - 1)</td>
</tr>
<tr>
<td>(R^S_t) Sovereign yield</td>
<td>Walrasian</td>
<td>Yes</td>
<td>No</td>
<td>(t_0 - t_4)</td>
<td>(t - 1)</td>
</tr>
<tr>
<td>(R^R_t) CLO yield</td>
<td>Walrasian</td>
<td>Yes</td>
<td>Yes</td>
<td>(t_2)</td>
<td>(t_2)</td>
</tr>
<tr>
<td>(R^D_t) Pension fund debt yield</td>
<td>Walrasian</td>
<td>Yes</td>
<td>Yes</td>
<td>(t_0)</td>
<td>(t_0)</td>
</tr>
<tr>
<td>(R^L_t) Bank loan rate</td>
<td>Banks</td>
<td>Yes</td>
<td>Yes</td>
<td>(t_1)</td>
<td>(t_1)</td>
</tr>
<tr>
<td>(Q_t) Capital</td>
<td>Firms</td>
<td>Yes</td>
<td>No</td>
<td>(t_1)</td>
<td>(t_1)</td>
</tr>
<tr>
<td>(W^C_t) Labor</td>
<td>Firms</td>
<td>No</td>
<td>No</td>
<td>(t_1)</td>
<td>(t_1)</td>
</tr>
<tr>
<td>(P^C_t) Consumption good</td>
<td>Firms</td>
<td>Yes</td>
<td>No</td>
<td>(t_3)</td>
<td>(t_3)</td>
</tr>
</tbody>
</table>

Notes: With the exception of the market for consumption goods, each non-Walrasian market features price competition. The market for consumption goods is special because firms must commit to producing a certain quantity before demand is known. Once production has occurred, the strategic environment switches to price competition. See Kreps and Scheinkman (1983) for a more general discussion of such a setup.

\(^{27}\)See Appendix A for a more extensive discussion of the employed equilibrium concept.

\(^{28}\)Interestingly, Lucas and Sargent (1979) define market clearing in spirit of my proposed definition of general equilibrium: "One essential feature of equilibrium models is that all markets clear, or that observed prices and quantities are viewed as outcomes of decisions taken by individual firms and households". In contrast, market clearing is understood as a situation in which excess demand is zero, \(z(p^*) = 0\), herein.
Notice that pricing of all financial claims directly or indirectly occurs relative to the return of investing at the predetermined risk free rate. The latter is determined as the central bank commits to intervening in bond markets such that \( R_{FFR}^t \) matches some previously announced target \( R_T^t \) in equilibrium.\(^{29}\) I now turn to discussing in detail the proposed intratemporal sequence of events.

**Subperiod \( t_0 \): Leverage**

At the beginning of each period, households own an aggregate stock of liquid demand deposits \( w_{t_0}^L = d_{t_0}^{HH} \) in the amount of last period’s income and accumulated retirement savings in the form of illiquid pension fund debt and equity \( w_{t_0}^l = w_{t_0}^D + w_{t_0}^E \) with corresponding uncertain returns \( \hat{R}_D^t \) and \( \hat{R}_E^t \). Institutionally, \( w_{t_0}^E \) constitutes a direct claim on fund assets, whereas debt \( w_{t_0}^D \) is held via a broker. On the asset side, pension funds hold demand deposits \( d_{t_0}^{PF} \) and a fraction \( a_{t_0}^S \) of the government security \( S \) acquired in previous period’s primary market. As illustrated in Figure 9, all demand deposits are fully backed by central bank credit \( v_{t_0} \) at this stage. Bank deposits are labeled \( M1 \) because they are accepted as means of payment \( v_{t_0} \) at this stage. Bank deposits are labeled \( M1 \) because they are accepted as means of payment in exchange for goods and services in the economy. Conversely, pension fund IOUs are illiquid because they are not accepted as a means of payment and only pay off at retirement.

**Figure 9.** Aggregate balance sheet of the banking sector

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserves ( v_{t_0} ) (( M0 ))</td>
<td>HH deposits ( d_{t_0}^{HH} ) (( M1 ))</td>
</tr>
<tr>
<td></td>
<td>PF deposits ( d_{t_0}^{PF} ) (( M1 ))</td>
</tr>
</tbody>
</table>

Notes: At the beginning of each period, all bank deposits (\( M1 \)) are backed by central bank reserves (\( M0 \)). This is not always the case as banks can create \( M1 \) by originating commercial loans.

At this time, households may opt to repartition their individual retirement portfolio. For this, they rely on a projected benchmark asset return of \( \hat{R}_A^t \equiv \hat{R}_T^t + \bar{\nu}_t \) for each \( h \in J^{HH} \), where \( \bar{\nu}_t \) is the geometric mean of the historical risk premium sequence \( \{ \hat{R}_A^\tau - R_T^\tau \}_{\tau=0}^{t-1} \). In deciding on their portfolio composition, households face a fundamental tradeoff between maximizing projected

\(^{29}\)Since the only price that is nominally rigid in the current setup is consumption sector wage, conventional monetary policy is ineffective in the sense that it has real effects if and only if the nominal wage constraint binds. The nominal wage constraint binds during a crisis episode or if the central bank’s target rate announcements vary substantially from period to period.
returns and limiting risk exposure with the latter increasing in \( w^E_{ht0}/w^I_{ht0} \) because equity is a residual claim. I then assume that households maximize risk-adjusted, projected returns as follows:

\[
V_{0}^{HH} (i \in J^{PF}, j \in J^B, w^I_{ht0}) = \max_{w^E_{ht0}, w^D_{ht0}} w^E_{ht0} \bar{R}^D_{ht} + w^E_{ht0} \bar{R}^E_{ht} - \gamma^L_h w^I_{ht0} \frac{w^E_{ht0}}{w^I_{ht0}}^2
\]

\[
\text{s.t. } w^E_{ht0} = w^I_{ht0} - w^D_{ht0}
\]

\[
\bar{R}^D_{ht} = R^{D,b}_{jt}
\]

\[
\bar{R}^E_{ht} = \bar{R}^A_{ht} + (\bar{R}^A_{ht} - R^{D,p}_{it}) L_{it}
\]

where \( R^{D,b}_{jt} \) is the gross rate offered to each household by broker \( j \), \( R^{D,p}_{it} \) is the lowest gross rate charged to pension fund \( i \) across all brokers, and \( L_{it} \) pension fund \( i \)'s leverage. Idiosyncratic risk aversion is thus captured by and increasing in \( \gamma^L_h \in (0, \infty) \). On the demand side, pension funds maximize projected return on equity subject to an initial margin requirement \( \delta_I \) to be relaxed by a maintenance margin \( \delta_M \) in subsequent subperiods:\(^{30}\)

\[
V_{0}^{PF} (\bar{R}^A_{it}) = \max_{L_{it}} \bar{R}^A_{it} + (\bar{R}^A_{it} - R^{D}_{it}) L_{it}
\]

\[
\text{s.t. } L_{it} \leq \delta_I
\]

Since pension funds rely on the same asset return projection, \( \bar{R}^A_{it} \equiv R^F_t + \bar{v}_t \) for each \( i \in J^{PF} \), the household problem implies \( \bar{R}^A_{it} > R^D_{it} \) such that funds strictly prefer debt finance in equilibrium. We have a corner solution with maximal leverage \( L_{it} = \delta_I \), \( w^E_{it0} = \frac{1}{1+\delta_I} w^I_{it0} \), and \( w^D_{it0} = \frac{\delta_I}{1+\delta_I} w^I_{it0} \) for each \( i \). Institutionally, equilibrium emerges as the auctioneer matches supply and demand for debt and equity across the two competitive brokers and pension funds:\(^{31}\)

**Subperiod \( t_1 \): Production**

The corporate sector consists of a set of consumption goods producers \( J^C \) and a set of capital producers \( J^K \). In subperiod \( t_1 \), households choose to supply labor to a firm \( f \in J^F = J^C \cup J^K \cup J^U \), where \( J^U = \{0\} \) represents voluntary unemployment. Letting \( n_{hf,t} \) and \( x_{hf,t} \) denote a household’s firm-specific labor supply indicator and labor output, firm \( f \) produces the following sector-specific outputs,

---

\(^{30}\)See Fabozzi (2015) for a discussion of the institutional details when buying assets on margin.

\(^{31}\)See Appendix A for a discussion of the resulting equilibrium strategies.
$$y_C^{f_t} = \begin{cases} z_t^C [k_{ft}]^\alpha \left[ \mu^{LF} \int x_{hft} \, dh \right]^{1-\alpha} & \text{if } f \in J^C \\ 0 & \text{otherwise} \end{cases}$$

$$y_K^{f_t} = \begin{cases} z_t^K \left[ \mu^{LF} \int x_{hft} \, dh \right] & \text{if } f \in J^K \\ 0 & \text{otherwise} \end{cases}$$

Where $x_{hft} = 1 \Rightarrow x_{hjt}$ for each $j \neq f$ meaning that each worker can only be employed by one firm. As indicated, both goods are produced using technology $z$ and labor $x$, but consumption goods production also requires capital $k$. The labor market setup is akin to Weiss (1980) in that households feature a time-invariant distribution of individual skill $q_h \sim G_q$, but each household also chooses to exert a certain level of effort $e_{hft}$. A household’s individual contribution to firm $f$’s labor output is then given by $x_{hft} = n_{hft} q_h e_{hft}$.

Prior to making their labor supply decision, each household receives a vector of firm-specific wage offers $\{W_{hft}\}$. While consumption sector offers are nominally fixed, offers in capital sector are tied to the individual contribution $x_{hft}$ and thus take the form of a contract. This is because I assume that a worker’s individual contribution $x_{hft}$ is observed by the capital producers, but not the consumption goods producers.\textsuperscript{32}

When choosing to supply labor to the consumption goods sector, households may fail to get matched, in which case they receive the same unemployment benefit as the voluntarily unemployed. The reason why involuntary unemployment exists is that my economy’s labor market may exhibit excess labor supply \textit{in equilibrium}.\textsuperscript{33} The corresponding unemployment benefits are then calculated as a fraction $\lambda^U \in (0, 1)$ of the lowest current consumption sector wage $W_{hft}^U = \lambda^U \min_{f \in J^C} \{W_{ft}^C\}$. Since $W_{hft}^U$ is uniform across households and $|J^U| = 1$, it is notationally reduced to $W_t^U$.

Given the described wage offers and the uncertainty associated with applying for consumption sector jobs, I assume that labor supply is chosen to maximize the following maxmin or “worst case scenario” criterion.\textsuperscript{34}

\textsuperscript{32}This ensures that the lucrativeness of each worker’s outside option is increasing in their skill.

\textsuperscript{33}If this is the case, each applicant is hired with equal probability with $\hat{W}_{hft}(\omega)$ denoting the ex post realized wage given the event $\omega$.

\textsuperscript{34}While extreme, this formulation is motivated by the premise that “a bird in the hand is worth two in the bush” as real-world households likely cannot infer true probability (see Knight, 1921; Savage, 1954). The setup, which may
\[ V_1^{HH} \left( \{W_{hft}\}_{f \in J^F} \right) = \max_{f \in J^F} \left\{ \min_{\omega \in \Omega} \hat{W}_{hft}(\omega) - \zeta I[f = J^U] \right\} \]  

where \( \hat{W}_{hft}(\omega) \) denotes the ex post realized wage and choosing voluntary unemployment carries a fixed utility cost of \( \zeta > 0 \) for all \( h \).

Once hired, workers are assumed to exert effort based on what they perceive as “fair” compensation (see Akerlof and Yellen, 1990). In the capital goods sector, fairness is not questioned because remuneration is directly linked to individual performance. In contrast, wage cuts in the consumption goods sector may be perceived as unfair because there is no transparent link between individual performance and effective wages. I specifically assume that workers retaliate nominal wage cuts as follows:

\[ e_{hft} = I[W_{ft}^C \geq \delta W_{ft-1}^C] \]

which effectively prevents firms from lowering wages below \( \delta W_{ft-1}^C \) and thus gives rise to a nominal downward rigidity. Given the institutional setup described thus far, I now proceed by discussing in detail the objectives pursued by firms in subperiod \( t_1 \).

**Capital goods sector.** \(|J^K| > 1\) capital firms maximize contemporaneous monetary profits by choosing a capital rental price \( Q_{ft} \) and a wage function \( W_{ft}^K(x) \). We have,

\[ V_1^{FK} \left( z^K_t, k^K_{ft-1}, \{k^K_{jt-1}, P^K_{jt}, W^K_{ft} \}_{j \neq f} \right) = \max_{Q_{ft}, W^K_{ft}(x)} Q_{ft} \min \{ \bar{k}^D_{ft}(Q_{ft}), k^K_{ft} \} - \int_{j \neq f} n_{hft} W^K_{ft}(x_{ht}) \, dh \]

s.t. \[ k^K_{ft} = (1 - \delta^D)k^K_{ft-1} + y^K_{ft} \]
\[ y^K_{ft} = \mu LF \int z^K_t n_{hft} x_{ht} \, dh \]

where \( \bar{k}^D_{ft} \) denotes a specific firm’s residual demand given \( \{k^K_{jt}, P^K_{jt} \}_{j \neq f} \). Since competition operates along the price margin, the quantities \( \{y^K_{jt} \}_{j \neq f} \) may not be taken as given. For example, alternatively also be thought of as a manifestation of “certainty effect” described in Tversky and Kahneman’s (1986), is extreme because households may choose to work for a capital producer even though working for a consumption goods producer entails a higher wage almost surely. In particular, recall that each household is of zero measure such that even if each consumption goods sector applicant is hired with probability one, a countable set of (measure zero) households may wind up unemployed involuntarily.

---

35As indicated, \( \hat{W}_{hft}(\omega) \in \{W_{ft}^C, W_{ft}^U \} \) for \( f \in J^C \) and \( \hat{W}_{hft}(\omega) = W_{hft} \) for \( f \in J^K \cup J^U \).

36As such, my specification behaviorally motivates the ad hoc rigidity proposed in Schmitt-Grohé and Uribe (2016).

37If multiple firms choose the same rental price of capital, residual demand is allocated proportionally.
a capital producer can always attract more labor by offering a very lucrative wage contract. Equilibrium in this market is characterized by a uniform, market clearing price strategy \( Q_{ft} = Q_t \) and a wage contract \( W^K_{ft}(x) = W^K_t(x) = z^K_t Q_t x \) for each \( f \in J^K \).\(^{38}\) Holding effort \( e_{ht} \) constant, notice then that \( W^K_{ht} = W^K_t(x_{ht}) \) implies that a worker’s sectoral wage offer is increasing in their skill \( q_h \) because individual labor output is given by \( x_{ht} = e_{ht} q_h \).

**Consumption goods sector.** \(|J^C| > 1\) firms rent capital and hire labor to produce a homogeneous, non-durable consumption good \( C \). Since capital is rented, maximization and the implied equilibrium strategies are uniform across \( f \in J^C \) for each \( t \). Since firms have no \((M1)\) at the beginning of \( t_1 \), capital rentals and worker salaries are financed with a bank loan \( l_{ft} \) at the interest rate \( R^L_{ft} \). Since consumption demand is a function of the uncertain taste shock \( \xi_t \sim G_{\xi_t|\xi_{t-1}} \) at the time of production, \( R^L_{ft} \) features a credit spread accounting for the case in which realized sales are insufficient to cover the face value of the loan. For purposes of realism and computation, it is further assumed that firms and banks locally approximate actual demand \( \hat{P}^C_{t_1}(y^C_t; \xi_t) \) with a function of the isoelastic type \( \hat{P}^C_{t_1}(y^C_t; \xi_t) \equiv \chi^k_t(\xi_t)[y^C_t]^{-\chi^r_t(\xi_t)}.\(^{39}\)

**Figure 10.** Exact and approximate aggregate demand for two realizations of the taste shock \( \xi_t \in \Xi \)

---

\(^{38}\)For a derivation of the firms’ equilibrium strategies, refer to Appendix A.

\(^{39}\)The approximation parameters \( \chi^k_{f_1}(\xi_t), \chi^r_{f_1}(\xi_t) \) are recovered numerically via a local symmetric difference quotient near last period’s aggregate output.
Figure 10 illustrates that $\tilde{P}^C_{t+1}(y^C_t; \xi_t)$ serves as a very accurate approximation of $P^C_t(y^C_t; \xi_t)$ with the latter indeed being approximately isoelastic, at least over the displayed domain. After observing the realization of the taste shock in subperiod $t_1$, firms compete for customers via price which implies market clearing in equilibrium. Ex ante, firms are then assumed to maximize expected profits as follows.

$$V^{FC}_t(z^C_t, W^C_{ft-1}; R^L_{ft}, Q_t) = \max_{l_{ft}, n^C_{ft}, W^C_{ft}} \mathbb{E}_{t_1}[\tilde{\Pi}_{ft}]$$

subject to:

$$\tilde{\Pi}_{ft} \equiv y^C_{ft} k^C_{ft}(\xi_t) \left[ \sum_{j \in C} y^C_{ft} \right]^{-\lambda}_{ft}(\xi_t) - l_{ft} R^L_{ft}$$

$$y^C_{ft} = z^C_t [k^C_{ft}]^\alpha \mu^{LF} \int n_{hf} q_{ht} c_{ht} \, dh$$

$$k^C_{ft} = (l_{ft} - \mu^{LF} n^C_{ft} W^C_{ft} - T^C_t) / Q_t$$

$$n^C_{ft} \leq \int n_{hf} \, dh$$

$$q^C_{ft} = f(W^C_{ft})$$

$$c^C_{ft} = 1[W^C_{ft} \geq \delta W^C_{ft-1}]$$

where the expectation is taken with respect to $G_{\xi_t|\xi_{t-1}}$ and the fact that the function $f$, to be derived shortly, is weakly increasing in $W^C_{ft}$ may generate an additional downward rigidity.

To understand the properties and the role of the function $f$, recall that each worker’s capital sector wage offer $W^K_{ht} = z^K_t Q_t q_{ht}$ is increasing in their skill $q_{ht}$, whereas — since optimization is symmetric across consumption goods producers — $W^C_t$ is uniform across across households. Given the utility specification in (1), a consumption sector wage $W^C_t$ thus attracts all households satisfying $q_{ht} \leq \delta^W W^C_t / z^K_t Q_t$. Holding fixed technology and prices, there then exists a marginal worker $h^*$ with skill $q^*_t \equiv \delta^W W^C_t / z^K_t Q_t$. We then partition the set of households as follows: Since each worker can only apply for one position at a time, households with $q_{ht} > q^*_t$ choose to supply labor to the capital goods sector, whereas households with $q_{ht} \leq q^*_t$ apply for consumption sector jobs. Average sectoral productivity is then plotted in Figure 11a and is given by,

40The realization of $P^C_t(y^C_t; \xi_t)$ in turn determines the return of the bank’s loan $\hat{R}^L_{ft} = \min\{R^L_{ft}, y^C_{ft} \tilde{P}^C_{t+1} / l_{ft}\}$.

41Firms understand that markets will clear ex post, but take other firms’ announced output as given in Cournot fashion ex ante. They thus implicitly assume that labor is not scarce in aggregate and thus do not recognize the possibility of affecting another firm’s output by poaching its workers (by offering an infinitesimally higher wage). If this were allowed, the competitive nature of Bertrand competition would have to induce $\mathbb{E}_{t_1}[\tilde{\Pi}_{ft}] = 0$. While potentially more realistic and certainly very interesting, the corresponding analysis is, like price commitment, beyond the scope of this paper and thus left for future work. For a detailed discussion of the equilibrium strategies, refer to Appendix A.
\[ q_t^K = \mathbb{E}[q | q > q_t^*] \]
\[ q_t^C = \mathbb{E}[q | q \leq q_t^*] \]

such that average labor productivity in the consumption goods sector \( q_t^C = \mathbb{E}[q | q \leq \delta^U W_t^C / z_t^K Q_t] \) is increasing in the wage offer \( W_t^C \). This is because increasing the wage attracts more productive workers, which raises average labor productivity. Conversely, lowering wage offers always induces the highest skilled workers to quit first. Depending on the cross-sectional distribution of skill, adverse selection may then induce a lower wage threshold below which firms will never optimally make an offer. In particular, if the skill distribution \( G_q \) induces a unique maximizer \( W_t^C \) of

\[
\mathbb{E} \left[ q | q < q^* \right] / q^* 
\]

offering any wage \( W_{t_i}^C < W_t^C \) is strictly dominated by the strategy of offering the lower bound \( W_t^C \). This is because below the threshold, labor costs are effectively decreasing in the wage. Such a situation is shown in Figure 9, where the initial, local concavity of \( \mathbb{E}[q | q < q^*] \) induces a unique, interior maximum of \( \mathbb{E}[q | q < q^*] / q^* \). In such a setting, optimality implies that firms will never offer wages below the threshold.\footnote{See Appendix A or Weiss (1980) for a proof.}

**Figure 11.** Pairing heterogenous skill with asymmetric information

Notes: Figure 11a depicts average sectoral labor productivity for a rectified Gaussian distribution \( G_q \), whereas Figure 11b illustrates the implied effective labor productivity per unit of wages paid (recall \( q_t^* = \delta^U W_t^C / z_t^K Q_t \), where everything but \( W_t^C \) is taken as given by the firm). By definition, the conditional expectation \( \mathbb{E}[q | q < q^*] \) is increasing in \( q^* \), which implies the following adverse selection mechanism from the point of view of firms. Lowering wage offers always induces the highest skilled workers to quit first, which depresses average labor productivity. This is not enough to generate the additional wage rigidity, for which I specifically require that (2) have an interior maximum such as the one highlighted in Figure 11 at \( x \approx 0.87 \). As illustrated, this desired property is satisfied by the chosen rectified Gaussian distribution because its mass at zero induces \( \mathbb{E}[q | q < q^*] \) to be locally concave.
Banks. Rather than disbursing commercial loans in reserves, banks artificially create previously nonexistent deposits (see Werner, 2014). This is acceptable to firms because deposits are accepted as means of payment irrespective of their backing.

Figure 12. $M_1$ creation by banks

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserves ($M_0$)</td>
<td>Deposits ($M_1$)</td>
</tr>
<tr>
<td>+ Commercial loan</td>
<td>+ Firm deposit ($M_1$)</td>
</tr>
</tbody>
</table>

Notes: Commercial loans are not disbursed in reserves, but firms are credited with previously nonexistent deposits. As long as the reserve requirement is not binding, banks can thus instantly and artificially rescale the economy’s numéraire $M_1$.

Assuming a reserve requirement of $\lambda_{RR}$, the banking sector is subject to the following aggregate constraint,

$$\lambda_{RR} \left[ d_{t0}^{HH} + d_{t0}^{PF} + d_{t1}^{F} \right] \leq v_{t0} \quad (3)$$

which implies that the central bank can curb lending by reducing the amount of aggregate reserves (see Blinder and Stiglitz, 1983; Bernanke and Blinder, 1992). Conversely, as impressively illustrated by the fact that the money multiplier $M_1/M_0$ was below 1 for virtually an entire decade between 2008 and 2018, boosting reserves in an effort to expand lending likely constitutes a doomed attempt to “push on a string”. 43 This is because banks need not use their reserves to extend loans. 44

In particular, rather than issuing the maximum possible amount of credit, banks can purchase risk free government bonds or lend to other banks instead,

$$d_{t1}^{F} \leq \frac{1 - \lambda_{RR}}{\lambda_{RR}} \left[ v_{t0} + \sum_{JB} a_{bt}^{B} - \sum_{JB} a_{bt}^{S} P_{t}^{S} \right] \quad (4)$$

where $a_{bt}^{S}$ denote the quantity of $S$ held by bank $b$ and borrowed reserves must be zero in

---

43The pushing-on-a-string phrase dates back to the depression era during which it was used — by the incumbent Fed Chair Eccles (Wood, 2005) among others — to describe the impotence of contemporary monetary policy.

44While excess funds were lent overnight to other depository institutions prior to 2008, they now earn IOER when held with the Fed. Unsurprisingly, this has induced banks to hold vast amounts of excess reserves.
aggregate. Assuming that banks approximate future demand just like firms, or $\tilde{P}_{Ct}^b = \tilde{P}_{Ct}^C$ for each $b \in J^B$, I propose the following bank objective.

$$V^B_1(v_{bt0}) = \max_{R^L_{bft}, a^L_{bft}, a^S_{bt}, v_{bt1}} \min_{\xi_t \in \Xi} \left\{ \left( \sum_{j \in J^C} \left[ l_{bft} - a^L_{bft} P^L_{bft} \right] \tilde{R}_{bft}^L(\xi_t) + a^S_{bt} V^S_t + v_{bt1} \right) - d^T_{t1} - v^B_{bt1} R^FFR_t \right\}$$

s.t. $d^F_{bft1} = l_{bft}$

$$d^T_{t1} = d^{HH}_{bt0} + d^{PF}_{bt0} + \sum_{j \in J^C} d^F_{bft1} - a^L_{bft} P^L_{bft}$$

$$d^T_{t1} \leq \frac{1}{\lambda_{RR}} v_{bt1}$$

$$v_{bt1} = v_{bt0} - a^S_{bt1} P^S_t + v^B_{bt1}$$

$$P^L_{bft} = l_{bft} \tilde{R}_{bft}^L / R^S_t$$

$$\tilde{R}_{bft}^L(\xi_t) \equiv \min \left\{ R^L_{bft}, \frac{y_{ft}^C \tilde{P}^C_t(\xi_t)}{\sum_{j \in J^B} l_{bft}} \right\}$$

To understand the bank’s problem, first notice that retaining a commercial loan on balance sheet entails the risk of generating a (marginal) profit short of the risk free rate. Given their maxmin risk preferences, banks will thus never opt to keep loans on balance sheet. Instead, banks follow an originate-to-distribute model in which each new loan is sold in its entirety ($a^L_{bft} = 1$).\textsuperscript{45} Loan sales are contractually set up by an auxiliary investment management company whose reservation price for the loan $P^L_{bft}$ is taken as given.\textsuperscript{46} Each loan’s interest rate is thus not determined by the fundamental risk preference of the issuing bank, but by the prevailing secondary market price $P^L_{bft}$.

In effect, Bertrand competition among banks then leads to competitive loan pricing $l_{bft} = P^L_{bft}$ and simultaneously ties the FFR to the exogenously set risk free interest rate $R^FFR_t = R^S_t$. The latter mechanism is known and exploited by the central bank in its setting of interest rates. As illustrated in Figure 13, it approximately holds in the data.

\textsuperscript{45}See Gorton and Pennacchi (1995) for a historical account of the originate-to-distribute model and the typical moral hazard concerns associated with it. Brunnermeier (2009) argues that originate-to-distribute led to a significant deterioration of lending standards in the early 2000s and thus played a major role in the creation of the “housing bubble” prior to the Great Recession.

\textsuperscript{46}It is assumed here that all financial market agents probabilistically assess the loans like a bank.
Figure 13. FFR vs. one-year treasury yield

Notes: Figure 13 illustrates the intimate empirical relationship between the Federal Funds Rate (FFR) and the one-year Treasury yield. In my theory, FFR and the one-period sovereign bond yield must be equivalent in equilibrium. This because investing in sovereign bonds yields the announced risk free rate such that banks are not willing to lend out their reserves returns at rates below it. At the same time, they are also not willing to borrow reserves at rates higher than the risk free rate because it would not be profitable. Both displayed series were taken from FRED.

Pension funds. Recall that pension funds began the period with an initial level of demand deposits \( d_{it_0}^{PF} \) and government securities \( a_{it_0}^{S} \). Having chosen their capital structure in subperiod \( t_0 \), they now proceed by maximizing assets returns subject to a liquidity constraint imposed by the government regulator.

\[
V_{it}^{PF}(d_{it_0}^{PF}, a_{it_0}^{S}, w_{it_0}^D) = \max_{d_{it_1}^{PF}, a_{it_1}^{S}} d_{it_1}^{PF} + a_{it_1}^{S} V_{it}^{S} \\
\text{s.t. } d_{it_1}^{PF} + a_{it_1}^{S} P_{it_1}^{S} \leq d_{it_0}^{PF} + a_{it_0}^{S} P_{it_1}^{S} \tag{5} \\
d_{it_1}^{PF} \geq \delta_L (d_{it_1}^{PF} + a_{it_1}^{S} P_{it_1}^{S}) \tag{6}
\]

Since holding government bonds is more lucrative than holding money, because \( V_{it}^{S} > P_{it}^{S} \), equation (5) binds in equilibrium.

Subperiod \( t_2 \): Securitization

The financial sector allows households to indirectly hold claims in firms via pension funds and a collateralized loan obligation \( R \). The latter is originated as the investment management company organizes all loan contracts in form a special purpose vehicle, whose shares are then sold to pension funds at \( t_2 \). Letting each commercial loan’s face value, true ex post payoff, and projected payoff be denoted by \( V_{bft} \equiv l_{bft} R_{bft}^{L} \), \( \hat{V}_{bft} \equiv l_{bft} \hat{R}_{bft}^{L} \), and \( \tilde{V}_{bft} \equiv l_{bft} \tilde{R}_{bft}^{L} \), the CLO’s face value, ex post payoff, and projected payoff are given by,
\[ V_t^R = \sum_{jB} \sum_{jc} V_{bft} \]
\[ \hat{V}_t^R(\xi_t) = \sum_{jB} \sum_{jc} \hat{V}_{bft}(\xi_t) \]
\[ \tilde{V}_t^R(\xi_t) = \hat{V}_t^R(\xi_t) \]

for each \( i \in J^B \cup J^{PF} \). Again, we have \( \hat{V}_t^R(\xi_t) \neq \hat{V}_t^R(\xi_t) \) because of the isoelastic price approximation.\(^{47}\)

Institutionally, the sale of \( R \) is conducted by the Walrasian auctioneer as pension fund demand is determined by,
\[
V_{2}^{PF}(d_{it1}^{PF}, a_{it1}^{S}, w_{it0}^{D}) = \max_{d_{it2}^{PF}, a_{it2}^{S}, a_{it2}^{R}} d_{it2}^{PF} + a_{it2}^{S} V_{t}^{S} + a_{it2}^{R} \mathbb{E}_{t2} \left[ \hat{V}_{it}^{R} \right] \\
\text{s.t.} \quad d_{it2}^{PF} + a_{it2}^{S} P_{it}^{S} + a_{it2}^{R} P_{it}^{R} \leq d_{it1}^{PF} + a_{it1}^{S} P_{it}^{S} \quad (7)
\]
\[
d_{it2}^{PF} \geq \delta L w_{it0}^{D} \quad (8)
\]

As long as funds are sufficiently liquid (including all \( S \) holdings), equilibrium must satisfy,
\[
P_{t2}^{R} = \frac{\mathbb{E}_{t2}[\hat{V}_{t}^{R}]}{R_{it}^{S}} \quad (9)
\]

where \( R_{it}^{S} \equiv V_{t}^{S}/P_{it}^{S} \). Equation (8) thus establishes (subjective) certainty equivalence between the lotteries associated with holding \( R \) and holding \( S \). Since the investment management company appraise loans based on their value as part of the CLO, we have,
\[
P_{bft}^{L} = \frac{l_{bft}}{\sum_{jB} \sum_{jc} l_{bft}} P_{t2}^{R} \quad (10)
\]

such that, by (9), \( P_{bft}^{L} \) is decreasing in \( R_{it}^{S} \) and thus \( R_{it}^{T} \). This relationship simply captures that monetary policy operates along Bernanke’s “cost-of-capital” channel, which postulates that firms’ cost of finance is higher when monetary policy is tight (2007). Since all debt is rolled over each period, neither firm nor bank net worth play an amplifying role in the model.

**Subperiod \( t_3 \): Marking-to-market**

At the beginning of \( t_3 \), financial markets observe a signal \( \xi_t' \sim G_{\xi_t'|\xi_t} \) with probability \( \pi_s \). If

\(^{47}\)We thus have \( \mathbb{E}_{t2}[\hat{V}_{it}^{R}] = \int_{\Xi} \hat{V}_{it}^{R}(\xi_t) dG_{\xi_t'|\xi_{t-1}} \) for each \( i \in J^{PF} \).
observed, the signal is processed to generate the newly relevant conditional distribution $G_{\xi_t | \xi_t', \xi_{t-1}}$ using Bayes’ theorem.\footnote{We have, \[ \Pr(\xi_t | \xi_t', \xi_{t-1}) = \frac{\Pr(\xi_t' | \xi_t, \xi_{t-1}) \Pr(\xi_t | \xi_{t-1})}{\Pr(\xi_t')} \] such that, as long as $G_{\xi_t' | \xi_t}$ and $G_{\xi_t | \xi_t', \xi_{t-1}}$ are known, it is straightforward to find $G_{\xi_t | \xi_t', \xi_{t-1}}$.} As projected demand is updated, funds revise their portfolios accordingly,

\[
V_{3}^{PF}(d_{it_3}^{PF}, a_{it_2}^{S}, w_{it_0}^{D}) = \max_{a_{it_3}^{R}, d_{it_3}^{PF}} a_{it_3}^{S} V_{t}^{S} + a_{it_3}^{R} \mathbb{E}_{t_3}[\bar{V}_{t}^{R}] + d_{it_3}^{PF}
\]

\[\text{s.t. } d_{it_3}^{PF} \geq \delta_{L} w_{it_0}^{D} + a_{it_3}^{S} p_{t}^{S} + a_{it_3}^{R} p_{t}^{R} \geq \delta_{M} w_{it_0}^{D}\]

where $\mathbb{E}_{t_3}[\bar{V}_{t}^{R}] = \int_{\Xi} \bar{V}_{t}^{R}(\xi_t) dG(\xi_t | \xi_t', \xi_{t-1})$. Because the liquidity constraint still binds across all funds, no purchases or sales of $R$ take place. The new equilibrium satisfies,

\[
P_{3}^{R} = \frac{\mathbb{E}_{t_3}[\bar{V}_{t}^{R}]}{R_{t}^{S}}
\]

such that, if a signal was observed, information-motivated trading has yielded a new risk-adjusted yield $R_{it_3}^{E} \equiv V_{t}^{R} / P_{3}^{R}$ with the risk free rate remaining unchanged. Each fund’s equity $w_{it_3}^{E}$ is recalculated as a residual of assets net of debt $w_{it_3}^{I} - w_{it_0}^{D}$. If the maintenance margin requirement $\delta_{M} w_{it_3}^{E} \geq w_{it_0}^{D}$ is violated, the broker issues a \textit{margin call} $MC_{it} \in \{0, 1\}$ demanding fund $i$ bring its account up to the minimum maintenance level by repurchasing part of its debt using demand deposits.

\[
MC_{it} = 1(\delta_{M} w_{it_3}^{E} < w_{it_0}^{D})
\]

\[
= 1(w_{it_3}^{I} < \frac{\delta_{I}(1 + \delta_{M})}{\delta_{M}(1 + \delta_{I})} w_{it_0}^{I})
\]

\[
= 1(w_{it_3}^{I} < \delta_{MC} w_{it_0}^{I})
\]

with a corresponding liquidity demand of $\Delta_{it}^{MC} = w_{it_0}^{D} - \delta_{M} w_{it_3}^{E}$. Facing a margin call, a fund’s otherwise prevailing liquidity constraint $d_{it_3}^{PF} \geq \delta_{L} w_{it_0}^{D}$ is lifted such that it may consolidate its position with the broker. However, since $d_{it_3}^{PF}$ assumes non-negative values only, the fund may find itself in a position in which it must procure fresh liquidity either by selling part of its portfolio or the issuance of new debt.\footnote{See Brunnermeier and Pedersen (2009) for an extensive account of funding vs. market liquidity.} If issuing new debt is impossible, brokers are assumed to proceed by \textit{selling out} part of the fund’s assets.

\[
\text{30}
\]
\[
SO_{it} = 1(d_{it2}^{PF} < \Delta_{it}^{MC}) = 1 \left( w_{it3}^I < \delta_I (1 + \delta_M - \delta_L) w_{it0}^I \right) = 1 \left( w_{it3}^I < \delta_{SO} w_{it0}^I \right)
\]

We must have \(\delta_{SO} < \delta_{MC}\), which implies that all sellouts are preceded by margin calls, but not all margin calls are followed by a sellout. Letting \(\Delta^G_{it} = \Delta_{it}^{MC} - d_{it2}^{PF}\) denote each fund’s liquidity gap, notice that the sale of any convex combination \([(\Delta^S_{it}, 0), (0, \Delta^R_{it})]_\lambda\) with \(\Delta^S_{it} = \Delta^G_{it}/P_{it3}^S\), \(\Delta^R_{it} = \Delta^G_{it}/P_{it3}^R\) and \(\lambda \in [0, 1]\) would be sufficient to make the broker whole. However, since sellouts are symmetric across funds, no fund is capable of offering \(M1\) in exchange for either security. The only remaining option is to sell assets to the central bank. Absent any unconventional monetary policy, the central bank only purchases the risk free security at the previously announced target. Therefore, fund \(i\) can satisfy its broker’s margin call if and only if the sale of \(a^S_{it2}\) is sufficient to cover \(\Delta^G_{it}\).\(^{50}\) Otherwise, the broker proceeds to sell out \(R\) in a fire sale.

\[
FS_{it} = 1(a^S_{it2} P_{it3}^S < \Delta^G_{it}) = 1 \left( w_{it3}^I < \delta_{FS} w_{it0}^I \right)
\]

Again, we must have \(\delta_{FS} < \delta_{SO}\) such that each fire sale is preceded by a sellout, but not each sellout is followed by a fire sale. In context of a fire sale, since there exists no private buyer for the risky security, the Walrasian auctioneer fails to locate an equilibrium price because the law of demand fails to hold: Lowering the security’s price induces lower demand because margin calls increase. At this stage, I proceed by assuming that the monetary authority intervenes and provides liquidity in some fashion. In the benchmark case, policy makers agree to buy up the entire excess supply at a predetermined haircut \(\delta_{HC}\) below the fundamental price \(P_{it3}^R\). Alternatively, the central bank could also enter into repurchase agreements, install an emergency lending facility, and/or lower the interest rate target.

\(^{50}\)Since \(S\) remains liquid while \(R\) does not, asset composition of each fund’s portfolio is crucial at this stage.
Subperiod $t_4$: Consumption

Recall that since households live for $T^L$ periods, there are $T^L$ overlapping generations at each time. Over the first $T^R$ periods of life, households belong to the working-age population and supply labor to firms. While in the labor force, workers accumulate retirement savings by investing part of each period’s income in illiquid financial claims. At retirement, the stock of previously accumulated financial wealth is liquidated and deposited in the corresponding household’s bank account. During retirement, households receive a pension from the government and draw down their accumulated savings until the age of $T^L$, at which point they are replaced by a new household.

Since accumulated savings are illiquid until retirement, households cannot boost next period’s consumption by saving more this period. Instead, the relevant benefit associated with the cost of decreasing consumption today $c_{ht}$ is given by a corresponding increase in projected retirement savings $\tilde{w}^P_{ht}$. The latter generate utility because households find retirement consumption too difficult to assess probabilistically, but understand that it is strictly increasing in accumulated savings.

As indicated by its designation, the projection $\tilde{w}^P_{ht}$ serves as a household’s best estimate of the effectively available funds at the time of retirement. Inspired by the Bureau of Labor Statistics’ (BoL) guidelines on how to save for retirement, $\tilde{w}^P_{ht}$ is derived via cumulative compounding with a benchmark return of $\tilde{R}_t^{A}$. Specifically, if $\tau^R_{ht}$ denotes household $h$’s remaining number of periods in the labor force at time $t$, I assume,

$$\tilde{w}^P_{ht} = w_{ht0}^I [\tilde{R}_t^{A}]^{\tau^R_{ht}} + s_{ht} w_{ht0}^L \sum_{i=1}^{\tau^R_{ht}-1} [\tilde{R}_t^{A}]^{i-1}$$

where $w_{ht0}^I$ is the beginning of period stock of illiquid retirement balances, $w_{ht0}^L$ are current liquid balances to be split between consumption and saving, and $s_{ht}$ is the chosen savings fraction. In words, households base their retirement balance projection on the assumption that until retirement, they will save the exact same nominal amount each period and that all savings will generate a return of $\tilde{R}_t^{A}$.

---

51 This assumption is by no means necessary, but it serves as a convenient way of emphasizing my main point, which is that the primary reason why households save is to accumulate wealth for retirement, not to shift consumption from today to tomorrow. An alternative specification would be to endogenize retirement with household utility being decreasing in the projected retirement age.

52 For purposes of illustration, consider the following concrete example taken from the US DoL publication “Top 10 Ways to Prepare for Retirement” (see Appendix C): Suppose you saved $5,500 each year until retirement in 35 years and your money earned 7% annually. In that case, your projected retirement balance would be $760,303.
For purposes of evaluation, households assess $\tilde{w}_t^P$ relative to a predetermined retirement goal $w_{ht}^G$ that aims to capture the idea that accumulated savings substitute for labor income during retirement,

$$w_{ht}^G = (T^L - T^R)(1 - \lambda_R)\tilde{W}_{ht-1}$$

where $\lambda_t W_{t-1}$ is the projected per-period retirement benefit received from the government.\footnote{As discussed shortly, the true retirement benefit is calculated as a fraction of the current going wage $W_{t-1}$ on a period-to-period basis.}

To generate the desired aggregate demand uncertainty, utility derived from retirement savings is subject to an aggregate shock $\xi_t$. Following the realization of $\xi_t$, firms compete to sell their existing inventories via price. Since production costs are sunk, equilibrium is given by a uniform price strategy $P_t^C$ that clears the market.\footnote{Since firms need not commit to a price strategy prior to the observation of $\xi_t$, the entire taste shock is absorbed via price. This leads to unreasonably volatile rates of inflation. Introducing price commitment thus constitutes a natural next step in improving the model. See Appendix A for a discussion of equilibrium.}

$$\int_{JHH} c_{ht} = \sum_{jC} y_{ft}^C$$

Finally, I assume that households only care about contemporaneous consumption and savings once retired. We have,

$$V_{4HH}(w_{ht0}^L, w_{ht0}^I) = \max_{s_{ht}} \begin{cases} u(c_{ht}; \gamma^c) + \xi_t v \left( \frac{\tilde{w}_{ht}^P}{w_{ht}^G}; \tilde{w}_{ht}^G \right) & \text{if in labor force} \\ u(c_{ht}; \gamma^c) + v \left( s_{ht} - \frac{\tau_{ht}}{\tau_{ht}^L + 1}; \gamma^s \right) & \text{if retired} \end{cases}$$

s.t. $c_{ht} = \frac{w_{ht0}^L (1 - s_{ht})}{P_t}$

$$\tilde{w}_{ht}^P = w_{ht0}^I [\tilde{R}_t^A]^R_{ht} + s_{ht} w_{ht0}^L \sum_{i=1}^{\tau_{ht}^R - 1} [\tilde{R}_t^A]^i$$

$$w_{ht}^G = (T^L - T^R)(1 - \lambda_R)\tilde{W}_{ht-1}$$

where, for purposes of this paper, $u$ and $v$ are specified as follows,

$$u(x; \alpha) = x^\alpha \quad \alpha \in (0, 1)$$

$$v(x; \alpha) = \frac{1}{\alpha} [1 - \exp(-\alpha x)] \quad \alpha > 1$$

Since $\lim_{x \to 0} u'(c; \alpha) \to \infty$, we must have $c_{ht} > 0 \ \forall ht$ in equilibrium. In effect, the above
utility specification induces the aggregate demand function depicted in Figure 10 while inducing a bell-shaped profile over lifetime wealth. Figure 14 illustrates the evolution of lifetime wealth across households who live for thirty periods of which they are retired for the last ten.

**Figure 14.** The life cycle of wealth in the cross-section

![Figure 14](image_url)

Notes: The above figure depicts the life cycle evolution of cross-sectional, beginning-of-period household wealth. Specifically, I report the mean as well as 68% and 95% confidence bands for a given age. During their working life, households accumulate wealth by saving their wage earnings and firm profits. In turn, following retirement, they proceed by financing part of their consumption by drawing down on accumulated savings. While labor income is zero during retirement, households still receive firm profits from their ownership of firms.

**Subperiod t5: Settlement**

As firms repay their loans, banks credit pension funds with additional demand deposits in the amount of $a_{it3}\hat{V}^R_t$ such that fund $i$’s realized asset, debt, and equity returns are given by,

\[
\hat{R}_t^A = w_{it4}^I/w_{it0}^I,
\]

\[
\hat{R}_t^D = \min\{w_{it4}^I, w_{it}^D \hat{R}_t^D\}/w_{it}^D,
\]

\[
\hat{R}_t^E = (w_{it4}^I - w_{it}^D \hat{R}_t^D)/w_{it}^E,
\]

where $w_{it4}^I = d_{it3}^{PF} + a_{it3}^S V_t^S + a_{it3}^R \hat{V}^R_t$. Next period’s funds, available for the purchase of the newly minted government bond, are then calculated as total assets net of expiring retirement balances,

\[
w_{it+10}^I = w_{it4}^I - \int_{\mu H} I(\tau_{hi}^R = 1) \left[w_{hit}^D \hat{R}_t^D + w_{hit}^E \hat{R}_t^E\right]
\]

Institutionally, the sovereign bond $S$ is rolled over as follows. First, in response to the prevailing
unemployment rate, the central bank issues next period’s interest rate target\textsuperscript{55},
\[
R_{t+1}^T = \max \{ \tilde{R}^T (1 - \bar{n}_t^u)^{\kappa_n}, 1 \}
\]

where $\bar{n}_t^u = n_t^u - n^u$ is unemployment net of the natural rate $n^u$. Given $R_{t+1}^T$, households revise their asset return projections\textsuperscript{56} and the treasury announces a new bond issue with face value $V_{t+1}^S$ to be paid back at the end of next period. Third, the new issue is sold in an auction open to all pension funds and the central bank.\textsuperscript{57} It is assumed that the central bank puts in a bid for the entire bond at $R_{t+1}^T$ thereby rendering the security risk free.\textsuperscript{58} The treasury thus knows it will at least generate revenues in the amount of $V_{t+1}^S/ R_{t+1}^T$ and pension funds know they must at least bid $R_{t+1}^T$ to receive any portion of the issue. It is further assumed that the treasury guarantees the central bank, in return for making its securities risk free, to supply sufficient liquidity for the latter to be able to implement its policy $V_{t+1}^S/ R_{t+1}^T = \sum_{JPF}(1 - \delta L)w_{it+10}^I$, which implies that $S$ is indeed priced by the central bank and not via Walrasian market clearing among funds. The final optimization problem of the period thus takes the following form,
\[
V_0^{PF}(a_{t+3}^S, a_{t+3}^{PF}; R_{t+1}^T) = \max_{R_{it+1}^S, d_{it+1}^{PF}} a_{it+10}^S V_{it+1}^S + d_{it+10}^{PF} R_{it+1}^T
\]
\[
s.t. \quad d_{it+10}^{SF} V_{it+1}^S/ R_{it+1}^T + d_{it+10}^{PF} = w_{it+10}^I
\]
\[
R_{it+1}^T \leq R_{t+1}^T
\]

where the equality of the last constraint neatly captures the central bank’s ability to set interest rates exogenously. In particular, notice that, since the market is sufficiently liquid, it is not optimal to submit a bid below $R_{t+1}^T$ such that each fund’s optimal strategy $(R^*, a^*)$ is given by $(R_{t+1}^T, (1 - \delta L)w_{it+10}^I R_{t+1}^T/V_{t+1}^S)$. To avoid immediate default, the new bond must also at least cover the government’s current liquidity gap.

\textsuperscript{55}Recall that the entire taste shock is absorbed along the price margin such that inflation targeting would not be contextually meaningful. Rather than camouflaging the fundamental pricing issue via a monetary targeting regime, the central bank is tasked to concern itself with unemployment only. Ex ante price commitment on the side of the firm constitutes a necessary next step because it would resolve this issue.

\textsuperscript{56}Households interpret the monetary policy change as a permanent level-shift and revise their asset return projection as follows: $\tilde{R}_{t+1}^A = R_{t+1}^T + \bar{\nu}_t$, where $\bar{\nu}_t$ denotes the average, additive risk premium as defined earlier.

\textsuperscript{57}The bond is allotted proportionally among the highest bidders, but pension funds are prioritized over the central bank.

\textsuperscript{58}While not in line with contemporary open market practice, the central bank engaging in the primary market is the only way to guarantee theoretical risklessness.
\[ V_{t+1}^S = \min \left\{ V_t^S + \left( X_t - T_t \right) - \Pi_t^{CB}, \sum_{JPP} (1 - \delta_L) w_{it+1}^I, R_{t+1}^T \right\} \]

where \( X_t \) denote government expenditures, \( T_t \) is total tax revenue, and \( \Pi_t^{CB} \) are returned central bank profits. Letting \( \tau_{ht}^R \) denote household \( h \)'s number of year until retirement, the evolution of household monetary wealth \( w_{hit+1}^L \) and illiquid wealth \( w_{hit+1}^F \) is given by,

\[
\begin{align*}
  w_{hit+1}^F &= 1(\tau_{ht}^R > 1) \left[ w_{hit_0}^L (1 - s_{ht}) + \sum_{JPP} w_{hit}^D \hat{R}_t^D + w_{hit}^E \hat{R}_t^E \right] \\
  w_{hit+1}^L &= (1 - \tau_I) \hat{W}_{ht} + w_{hit_0}^L (1 - s_{ht}) + \gamma_h^\beta \sum_{JFF} \hat{I}_{ft} + 1(\tau_{ht}^R = 1) \sum_{JPP} w_{hit}^D \hat{R}_t^D + w_{hit}^E \hat{R}_t^E
\end{align*}
\]

where \( \gamma_h^\beta \overset{i.i.d.}{\sim} G_\sigma \) denotes a household’s time-invariant ownership share of the firms. Retired (\( n_{hit}^R = 1 \)) and unemployed (\( n_{hit}^U = 1 \)) households receive a pre-tax wage of \( \lambda_R W_t^C \) and \( \lambda_U W_t^C \) respectively. Finally, government tax income and expenditures are given by \( T_t = \int_{JHH} \tau_I \hat{W}_{ht} \, dh \) and \( X_t = \int_{JHH} [n_{hit}^R + n_{hit}^U] \hat{W}_{ht} \, dh \). This concludes the period.
4 A typical fire sale episode

To conduct a quantitative analysis, the proposed economy must be parameterized. For this, I partition \( \theta \) into two subsets (\( \theta_1, \theta_2 \)). First, \( \theta_1 \) is calibrated\(^59\) with reference to the literature, regulatory rules, or to specific data statistics. Second, because the computational requirements render a likelihood approach (e.g. particle filtering) impractical, \( \theta_2 \) is estimated via simulated method of moments (SMM). In face of occasional, but rare crises, recovering relevant moments from the data is non-trivial. Primarily, this is because the number of observed crises in the employed dataset is so low that, even if observation horizons were sufficiently long to assert conditional convergence (crisis and non-crisis), the same is almost surely not true unconditionally. Instead of using recent US data to match the observed unconditional moments, via the observed crisis frequency, I thus proceed by borrowing the latter statistic from Barro’s (2006) treatment of “rare disasters” and subsequently targeting conditional moments only.

Now, before delving into the concrete moments to be matched, I would like to briefly pinpoint two principal issues arising from the Monte Carlo nature of SMM. First, since the duration of moment convergence crucially depends on the quality of state initialization, the system is initialized at the implied deterministic steady state for capital for each candidate \( \theta'_2 \). In an additional effort to improve finite sample performance, the selected sampling horizon \( \bar{T} \) is augmented with a burn-in period \( \bar{T} \). The total simulation horizon is then given by \( T = \bar{T} + \bar{T} \), where the first \( \bar{T} \) periods allow the economy to converge to its stationary distribution while time averages are only computed using data that is generated thereafter. Second, to ensure continuity of the loss function, the same sequence of random innovations must be recycled for all evaluations across \( \Theta_2 \). However, this implies that the resulting estimate \( \hat{\theta}_2 \) is not only a function of the data and the model, but also of a random quantity that is the sequence of randomly drawn shocks. Traditionally, the estimation uncertainty resulting from this dependence is addressed via bootstrapping — by drawing a swarm of different shock sequences — but computational expense prohibits such a procedure in this case.

\(^{59}\)Calibration is understood as any procedure in which \( \theta_1 \) is chosen, whereas estimation entails numerical (or analytic if possible) minimization of a predetermined, joint loss function of \( \theta_2 \) given the data \( \hat{\theta}_2 = \arg\min_{\theta_2} L(\theta_2; Y) \). Since identification is rarely verifiable across the entire space \( \Theta_2 \), even estimation requires careful ‘calibration’ of the initial guess \( \theta_{20} \).
Empirical targets

As indicated, rather than targeting unconditional statistics, I proceed by targeting moments that condition on whether or not the economy is in a state of crisis or not. The chosen statistics are summarized in Table 2 with $\kappa^c$ and $\kappa^{nc}$ denoting crisis and non-crisis statistics respectively,

<table>
<thead>
<tr>
<th>Panel A: Statistics targeted via calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_c$</td>
</tr>
<tr>
<td>$\kappa_1^c$</td>
</tr>
<tr>
<td>$\kappa_2^c$</td>
</tr>
<tr>
<td>$\kappa_3^c$</td>
</tr>
<tr>
<td>$\kappa_4^c$</td>
</tr>
<tr>
<td>$\kappa_2^{nc}$</td>
</tr>
<tr>
<td>$\kappa_3^{nc}$</td>
</tr>
<tr>
<td>$\kappa_4^{nc}$</td>
</tr>
<tr>
<td>$\kappa_5^{nc}$</td>
</tr>
<tr>
<td>$\kappa_6^{nc}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Statistics targeted via estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_7^{nc}$</td>
</tr>
<tr>
<td>$\kappa_8^{nc}$</td>
</tr>
</tbody>
</table>

As indicated, the targeted crisis frequency is taken from Barro’s treatment of “rare disasters” (2006) with a corresponding targeted employment recovery duration of ten years. Additional crisis targets include nominal wealth loss, calculated from the evolution of household net worth during the Great Recession, as well as the peak in output gap and unemployment recorded during the Great Recession. Data on unemployment and labor shares are obtained from the US Bureau of Labor Statistics, whereas all other data is taken from FRED. Finally, output is detrended via first differencing. Refer to Appendix B for a discussion of the data and the computation of each statistic.
Calibration

Table 3 summarizes the calibrated parameter values given the chosen time unit of one quarter.

<table>
<thead>
<tr>
<th>Panel A: Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Specific targets$^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_W$</td>
</tr>
<tr>
<td>$\delta_D$</td>
</tr>
<tr>
<td>$\lambda_U$</td>
</tr>
<tr>
<td>$\lambda_R$</td>
</tr>
<tr>
<td>$\gamma^c$</td>
</tr>
<tr>
<td>$\gamma^w$</td>
</tr>
<tr>
<td>$\gamma^r$</td>
</tr>
<tr>
<td>$\gamma^o$</td>
</tr>
<tr>
<td>$q$</td>
</tr>
<tr>
<td>$\bar{R}_T$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Regulatory parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{RR}$</td>
</tr>
<tr>
<td>$\lambda_M$</td>
</tr>
<tr>
<td>$\lambda_{IM}$</td>
</tr>
<tr>
<td>$\lambda_{MM}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: Metaparameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_L$</td>
</tr>
<tr>
<td>$T_R$</td>
</tr>
<tr>
<td>$</td>
</tr>
<tr>
<td>$\mu^{HH}$</td>
</tr>
<tr>
<td>$N^{HH}$</td>
</tr>
</tbody>
</table>

Notes: The set of calibrated parameters are partitioned into four subsets. First, the capital share of production in the consumption goods sector is set as is common in the literature. Second, nine parameters are set to target specific statistics from the data. The liquidity requirements are set based on observed practice, whereas the margin requirements deviate from observed practice in order to generate occasional fire sales. The metaparameters are chosen to scale the economy subject to the limitations imposed by computational constraints.

To match the initial labor market disruption, the nominal wage rigidity parameter is set to be $\lambda_W = 0.99$, which is in line with Schmitt-Grohé and Uribe (2016). Capital depreciation $\delta_D$ is set equal to 5% quarterly to generate the slow employment recovery and a deterministic steady state of unemployment of 5%. While the liquidity requirements are set equal to values that constitute
common practice, the margin requirements are calibrated independently. The unemployment benefit targets the prevailing level of unemployment, whereas retirement benefits are chosen in an ad hoc fashion. The distribution of skill $G_q$, risk preference $G_s$, and ownership $G_o$ are set as follows: Targeting a full employment labor share in the capital goods sector of 12%, $G_q$ was calibrated to induce a unique maximizer $x$ of $f_q(x) \equiv \mathbb{E}[q | q < x] / x$ at $x = 0.88$. In turn, $G_s$ generates the desired consumption share of income while ownership $G_o$ gives rise to the desired Gini coefficient for income. Finally, the number of overlapping generations was set to 30 to limit computational expense while the corresponding retirement age ensures a time-invariant working-age to total population ratio of $\frac{2}{3}$. The number of consumption sector firms is a natural metric for sectoral competition and was chosen to generate the desired credit spread on the commercial loans. The effectively simulated number of households was finite for computational reasons, and set equal to 50 per generation. The time-invariant measure of households $\mu^{HH}$ and the initial level of $M0$ only serve as tools to scale the economy, in real and nominal terms respectively, and were chosen to normalize initial output and the initial wage level to unity.

**Exogenous drivers**

I proceed by parameterizing the technology and taste shock processes via SMM. For this, I assume that capital technology evolves according to an AR(1) process with ergodic mean $\mu_k$, persistence parameter $\rho_k$, and a shock standard deviation of $\sigma_k$. For purposes of simplicity, the taste shock is reduced to a Markov chain of state size two, good and bad. In the good state $\xi^{good}$, households care less about retirement savings and aggregate demand is high. In the bad state $\xi^{bad}$, households care more about retirement savings and aggregate demand is low. I now estimate the

---

60 The initial and maintenance margin are set to be lower than the actual thresholds of 50% and 25% required under the Federal Reserve Board’s Regulation T because the corporate sector setup does not generate enough loan repayment volatility to induce a margin call at those values. This is primarily due to the fact that loans are not only identical ex ante, but even ex post. Introducing price commitment and/or a spatial reallocation of households along a circular city may mitigate this problem, but both are left for future work.

61 Since capital goods firms, pension funds and banks compete in Bertrand fashion, their corresponding number is irrelevant as long as it is greater than 1.

62 Given the vast dimensionality of the model, I do not solve for policy or price functions across the entire state space. However, finding such functions is not necessary because even though each optimization problem takes the form of a functional equation, the latter are never self-referential as they are in the canonical Bellman setup. In consequence, this means that equilibrium at each point in the state space can be found irrespectively of equilibrium at any other point. The model is thus simulated by recursively solving for equilibrium anew each period. Exploiting parallelization, each period then takes roughly 30 seconds to simulate across 24 cores on Vanderbilt’s computing cluster. Notice that the combination of household heterogeneity and the life cycle nature of the model introduces an additional layer of complexity because the number of simulated households is finite. To prevent generational cycles arising from time-invariant sources, it is crucial that parameterization be equivalent across all generations.
size of the bad shock $\xi^{\text{bad}}$ as well as the persistence $\rho_k$ and volatility $\sigma_k$ of the technology shock via SMM. To ensure continuity of the SMM loss function $\mathcal{L}$, the same sequence of random innovations is recycled for all evaluations across $\Theta_2$. Table 4 reports the corresponding results.

### TABLE 4. ESTIMATION

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}_k$</td>
<td>Capital technology persistence</td>
<td>0.905</td>
</tr>
<tr>
<td>$\hat{\sigma}_k$</td>
<td>Capital technology shock volatility</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Notes: The exogenous capital technology process is estimated via SMM. The parameters are just-identified as two moments are used to estimate two parameters. Since computation time is linear in the simulated horizon $T$, producing standard errors via Monte Carlo is infeasible.

The resulting capital technology process is then given by

$$
\xi^K_t = (1 - \hat{\rho}_k)\mu_k + \hat{\rho}_k \xi^K_{t-1} + \varepsilon^K_t, \quad \varepsilon^K_t \sim \mathcal{N}(0, \hat{\sigma}_k)
$$

The employed Markov chain is calibrated to induce an ergodic demand slump frequency of $\pi_{\text{recession}} = 0.3$. The signal frequency $\pi_{\xi' \in \Xi}$ and its accuracy — as given by the probability that the signal corresponds to the true state $\pi_{\xi' = \xi | \xi = i}$ — were set to match the proposed yearly frequency of rare disasters $\pi_c = 0.017$ in Barro (2006) while ensuring that a financial crisis emerges if and only if a bad state is followed by a bad signal.

$$
\pi_c = \left( \frac{\pi_{\xi' = \xi^{\text{bad}} | \xi = \xi^{\text{good}}} \pi_{\xi = \xi^{\text{good}} | \xi_{t-1} = \xi^{\text{bad}}} + \pi_{\xi' = \xi^{\text{bad}} | \xi = \xi^{\text{bad}}} \pi_{\xi = \xi^{\text{bad}} | \xi_{t-1} = \xi^{\text{bad}}}}{\pi_{\xi' = \xi^{\text{bad}} | \xi = \xi^{\text{good}}} \pi_{\xi = \xi^{\text{good}} | \xi_{t-1} = \xi^{\text{bad}}} + \pi_{\xi' = \xi^{\text{bad}} | \xi = \xi^{\text{bad}}} \pi_{\xi = \xi^{\text{bad}} | \xi_{t-1} = \xi^{\text{bad}}}} \right)
$$

The resulting process for the taste shock is summarized in Table 5.

### TABLE 5. THE TASTE SHOCK

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taste shock space</td>
<td>${0.01, 0.5}$</td>
</tr>
<tr>
<td>Markov chain</td>
<td>${0.86, 0.2}$</td>
</tr>
<tr>
<td>Signal frequency (i.i.d.)</td>
<td>0.05</td>
</tr>
<tr>
<td>Signal accuracy</td>
<td>${0.95, 0.997}$</td>
</tr>
</tbody>
</table>

A typical fire sale episode

Given the parameterized economy, I can now quantitatively examine the macroeconomic trans-
mission of a financial fire sale, an example of which is depicted in Figure 15.

**Figure 15.** A fire sale

![Graph showing Unemployment, Output, Investment, and Capital over time.](image)

Notes: Figure 15 depicts unemployment, output, investment, and the capital stock for a given realization of technology in a 20 year window surrounding a fire sale at time $t = 0$. On average, workers lose roughly 20% of their nominal lifetime savings during the fire sale. In absolute terms, older workers who are nearing retirement are hit the hardest because they have accumulated more wealth than their younger counterparts. At time $t = 1$, households respond to the nominal shock by curbing consumption demand in nominal terms. Consumption goods producers anticipate this and react, because they are unable to freely adjust wages given the ‘fair wage’ constraint, by reducing investment and labor demand. By the time the nominal wage constraint no longer binds, around five periods after the initial shock, the reduced capital stock depresses labor productivity such that labor demand only slowly reverts back to the original steady state level.

First, notice that unemployment immediately more than doubles following the fire sale at $t = 0$. This initial bump is caused by the behavioral constraint imposed by worker effort, whereas persistence arises from the asymmetric information surrounding idiosyncratic worker productivity. In particular, facing the initial slump in household demand, firms do not lower wages by more than $1 - \delta_W$ because otherwise workers would stop exerting any effort. This effect only lasts for a few quarters, after which excess unemployment persists because the lower level of capital depresses labor demand via labor productivity, an effect that is either mitigated or amplified by current technology. At that stage, firms do not lower wages to clear the market because of the selection issues arising from asymmetric information as discussed in Weiss (1980). Notice that the observed labor market recovery only takes thirty quarters and is thus quicker than the targeted window of ten years.

Importantly, notice that Figure 15 only depicts a single crisis for a given sequence of technology
shocks. We may then wonder what such a financial crisis looks like if the effects of technology are integrated out. For this, consider Figure 16 which depicts 100 simulated paths of my model economy for 100 paths of technology with the shaded areas representing 68% and 95% confidence bands.

**Figure 16.** A typical fire sale

![Figure 16](image_url)

Notes: Figure 16 was created by simulating 100 fire sales given 100 different paths for technology. The resulting mean orbit, as depicted by the black line, represents a typical financial fire sale in the sense that the temporary effects of technology have been integrated out. The two shaded areas represent 68% and 90% confidence bands for the respective values.

Figure 16 illustrates that the economy takes roughly ten years to revert back to its non-crisis ergodic benchmark. In particular, this implies that unemployment takes roughly ten years to recover *in expectation*. As indicated, the slow recovery is due to the loss in capital, which depresses labor productivity and takes roughly ten years to rebuild. Finally, consider Figure 17 in which I plot the estimated ergodic density for unemployment of my parameterized economy.

As desired, the resulting unemployment density features a substantial right tail including an additional mode with the corresponding statistics being recorded immediately following the fire sales. Crucially, further notice that the targeted crisis frequency borrowed from Barro (2006) is significantly lower than the observed frequency since the onset of the Great Moderation. The fact that the ‘crisis mode’ of the unconditional density depicted in Figure 17 is much less pronounced

---

64 In the limit, as the number of Monte Carlo simulations approaches infinity, the economy converges to its conditional (crisis vs. no-crisis) ergodic distribution in which case the confidence bands are smooth.
than its empirical counterpart in Figure 3 is thus by construction. Whether or not the model is satisfactory in rationalizing the data should thus primarily be judged by way of the two conditional, crisis and non-crisis, densities.

**Figure 17.** Estimated ergodic density of unemployment for the parameterized model economy

Notes: The above figure depicts three estimated ergodic densities for simulated unemployment akin to Figure 3. Qualitatively, the densities look as desired. There is a substantial right tail that is recorded during the economic downturns following the fire sales. The circumstance that the ‘crisis mode’ of the unconditional density is much less pronounced than its empirical counterpart in Figure 3 simply reflects the fact that the targeted crisis frequency was taken from Barro (2006), not from the data since 1987. For purposes of contrasting the two graphs, the relevant densities for purposes of comparing the model with the data are thus the two conditional densities.

**Implications for policy**

“How would behavior have differed had certain policies been different in specified ways?” (Lucas, 1977). The Great Depression and the Great Recession were both preceded by a financial crisis, but the monetary policy response varied greatly across the two episodes. In fact, Friedman and Schwartz (1963) view monetary policy as a primary root of the length and depth of the Great Depression, a perspective recently reevaluated by Romer and Romer (2013) and Christiano, Motto, and Rostagno (2003). Since my framework’s institutions were modeled to represent the US economy since the onset of the Great Moderation, the proposed theory does not provide a natural platform to quantitatively assess observed policies during the Great Depression. However, recall that deep downturns are caused by the widespread loss in retirement savings during the fire sales in my model.

**Insight.** Once a liquidity crisis is imminent, the success of crisis response may be measured by the degree to which it is able to contain nominal losses.

For example, to the extent that the Federal Reserve did not manage to prevent, or maybe even exacerbated, the nominal collapse during the Great Depression, my analysis suggests that policy...
was not successful at the time. Extending beyond the qualitative insight discussed so far, I now turn to conducting a counterfactual, quantitative policy evaluation for my parameterized model economy. For this, notice that the central bank can, as the ultimate creator of the numéraire, at any time decide to provide additional market and/or funding liquidity (see Brunnermeier and Pedersen, 2009). In particular, recall that the simulated path depicted in Figure 16 was generated by targeting an aggregate nominal wealth loss of twenty percent, a quantity that is not invariant to policy. We may then wonder how the economy would have fared under a variety of alternative, unconventional monetary policy regimes.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Channel</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>No unconventional policy</td>
<td>-</td>
<td>Pension fund bankruptcy</td>
</tr>
<tr>
<td>Cut risk free target rate</td>
<td>Equity</td>
<td>Price increase of S and R</td>
</tr>
<tr>
<td>Outright purchase of R</td>
<td>Liquidity</td>
<td>Reallocation of R</td>
</tr>
<tr>
<td>Repurchase agreement</td>
<td>Liquidity</td>
<td>Collateralized lending against R</td>
</tr>
<tr>
<td>Emergency lending</td>
<td>Liquidity</td>
<td>Collateralized lending against R</td>
</tr>
</tbody>
</table>

Notes: Table 6 lists a set of policy tools that the central bank has at its disposal during a financial fire sale. If policy makers decide to refrain from deploying unconventional policy, the Walrasian auctioneer drives the CLO price to zero and pension funds go bankrupt. Cutting the risk free interest immediately raises asset prices, which may already be enough to avert the crisis because it boosts pension fund equity. If not, the central bank may opt to purchase the CLO outright, to enter into a reverse repurchase agreement, or to lend to funds as a lender of last resort. Interestingly, any successful policy works by effectively circumventing the margin constraint imposed by the broker.

In face of a liquidity crisis, a natural first step to undertake for the central bank is to cut the previously announced interest rate target, thereby immediately boosting equity via asset prices. While this measure does not necessarily require any further action (e.g. asset purchases) on part of the central bank, it is likely insufficient to prevent a fire sale. If so, the central bank may provide additional liquidity via emergency lending or via outright purchases and repurchase agreements of non-governmental securities. In case of an outright purchase, the central bank agrees to take possession of the security ex ante, prior to the realization of its payout, whereas both repurchase agreements and emergency lending constitute a form of collateralized short-term lending, in which case the central bank assumes the risk of receiving the proceeds of the security ex post.

I start by assuming that, once a fire sale is imminent, the central bank proceeds by purchasing
the risky security outright in the open market. For this, consider Figure 18, which plots the macroeconomic transmission of a specific fire sale episode across different policy regimes. In particular, holding technology evolution constant, the different regimes are captured by a set of varying prices at which the central bank agrees to buy. Intuitively, the respective price thresholds metaphorically stand for the varying durations that the central bank waits as the Walrasian auctioneer iteratively lowers the resulting transaction price: The longer the central bank waits, the lower the price. As discussed previously, the severity of transmission uniquely depends on the degree to which nominal wealth collapses in my model. Therefore, as illustrated in Figure 18, more decisive central bank policy is mirrored by a less distressed real economy. In essence, once a crisis is imminent, more aggressive policies lead to less severe transmissions.

**Figure 18.** A fire sale across policy regimes

Notes: Figure 18 illustrates the macroeconomic transmission of the same financial crisis across a set of different policy regimes. In particular, each depicted series was generated using the same exact evolution of technology, but the price threshold at which the central bank starts absorbing the excess supply of the CLO during the fire sale at \( t = 0 \) was varied. The more aggressive the central bank’s intervention, the smaller the nominal impact of the fire sale, and the more muted the macroeconomic transmission. In fact, if the central bank is willing to buy at the CLO’s fundamental value, the looming transmission to the real sector can virtually be contained in its entirety.

Instituting repurchasing agreements or emergency lending, on the other hand, effectively amounts to an exchange of the pension funds’ counterparty: A portion of leverage is transferred from the broker, who imposes the maintenance margin, to the central bank which acts as a ‘lender of last resort’. Importantly, I assume that the central bank does not impose a maintenance margin, but
instead insists on receiving preferred status in the chronology of payouts. As such, collateralized lending institutionally prevents fire sales by effectively circumventing the margin constraint imposed by the broker.

**Insight.** Fire sales are not symptomatic of fundamentally deteriorating assets, but of contemporary financial market institutions.

But if liquidity crises are an institutional problem, it is unsurprising that circumventing the respective institutions, as suggested above, serves as an effective mitigation mechanism. However, notice that all of the previously proposed policy only addresses scenarios in which a crisis is already imminent. We may thus wonder what can be done to strengthen systemic resilience ex ante. In light of the gained insights, an intuitive policy recommendation would be to disallow maintenance margins altogether. Banning margin calls would effectively eliminate the source of all fire sales, but it would inevitably also raise the costs associated with leverage.

5 Discussion

Empirical evidence suggests that financial crises affect real activity via both supply and demand (Chodorow-Reich, 2014; Mian, Rao, and Sufi, 2013), but contemporary macroeconomic theory overwhelmingly emphasizes channels of aggregate supply. The core contribution of this paper then lies in the formalization of a dynamic, wealth effect driven aggregate demand channel. The proposed framework’s financial sector is modeled to be fragile in the sense that it gives rise to occasional financial crises in the form of fire sales. Fire sales are followed by a reallocation of nominal claims from liquidity-strapped pension funds to the central bank. In an effort to make up for lost retirement wealth, the working-age population reacts by cutting consumption expenditures, in anticipation of which firms scale back production via employment and investment. To generate the targeted, steep and persistent real effects following a nominal collapse, I require a nominal and a real wage rigidity. In addition to the proposed transmission mechanism, the framework offers a variety of insights regarding policy, the nominal accounting of $M1$, and macroeconomic methodology.

---

65I entirely abstract from tightening borrowing constraints faced by consumers, which likely also played a key role during the Great Recession (Mian, Rao, and Sufi, 2013).
On policy

In my model, conventional monetary policy only directly affects commercial lending by way of the traditional cost-of-capital channel (Bernanke, 2007). While the central banks’ interest rate announcements affect firms’ cost of production — and thus potentially aggregate supply\textsuperscript{66} — neither firm nor bank net worth play an amplifying role because all claims are rolled over each period. However, since households perceive monetary policy changes as permanent level-shifts, an increase in the policy rate also affects aggregate demand via intertemporal substitution. First, previously accumulated financial wealth is projected to generate more retirement wealth which depresses the incentives to save (wealth effect).\textsuperscript{67} Second, each unit of additional savings is projected to generate more retirement wealth, which encourages saving (substitution effect).

The term unconventional monetary policy, which is exclusively enacted at $t_3$ in my model, is used to describe any extraordinary action taken by the central bank to combat an imminent or ongoing financial crisis. For this, recall that when the maintenance margin constraint binds, the Walrasian auctioneer will drive the price of the CLO to zero unless the central bank intervenes. Naturally, the first option to consider is that the central bank continues to defend its interest rate target by purchasing as much of the sovereign bond from the pension funds until all brokers’ margin calls are satisfied. When this is insufficient, the central bank has an array of additional, unconventional tools at its disposal as previously illustrated in Table 6. While lowering the interest rate target boosts asset prices directly, all other tools boost market or funding liquidity. As discussed, once a liquidity crisis is imminent, more aggressive policies lead to less severe transmissions. One might be tempted to argue that prescribing more aggressive policy ex post likely encourages moral hazard by inducing more risky behavior ex ante. However, notice that the crises examined herein do not arise from undesirably risky behavior on part of the investors. Instead, crises emerge from an institutional constraint, the maintenance margin, that is designed to shield debt holders from losing their investments. In fact, notice that all policies that successfully mitigate or prevent the real transmission effectively do so by circumventing the maintenance margin. A natural way to reduce systemic risk ex ante would then be to disallow maintenance margins altogether, a measure

\textsuperscript{66}Conventional monetary policy has no real effects unless the nominal wage constraint binds as is the case during a crisis episode.

\textsuperscript{67}Since there exist no multi-period bonds, increasing the interest rate across periods does not depress the prices of any outstanding securities. The same is not true for unconventional policy enacted at $t_3$. 

48
that would most likely increase the cost of leverage.

**On the numéraire, nominal wealth, and the role of banks**

My economy’s unit of account is given by noninterest-bearing demand deposits issued by banks. Agents exhibit demand for such deposits for a number of reasons. Consumption goods producers take out loans from banks because bank deposits are the only means of payment accepted by both capital suppliers and workers. Capital producers only accept deposits as means of payment because worker salaries must be paid in deposits. Households accept and retain $M_1$ throughout the period because consumption goods must be purchased in exchange for deposits. Pension funds hold $M_1$ because they are required do so by law.

Money creation is carried out by the central bank ($M_0$) and by commercial banks ($M_1$). The latter can create money because their deposits are, irrespective of their balance sheet counterpart, accepted as a means of payment within the private sector. Specifically, banks can credit firms with previously nonexistent deposits without having to expense any of their reserves. While such demand deposits that are not backed by central bank credit may be viewed as artificial, because they are fictitiously “invented” by banks (Werner, 2014), such a system allows banks to promptly respond to money demand by scaling up and down the numéraire. Conversely, a “full-reserve” alternative would almost certainly require the central bank to alter, if not abandon the contemporary practice of targeting interest rates (SNB, 2018). Therefore, because banks are a key component of modern monetary systems in which the central bank tightly controls the nominal interest rates, but not monetary aggregates, the banking sector was modeled in the spirit of contemporary practice.

Since my framework does not feature long-term loans, it is impossible for a crisis to induce assets price dislocations across periods. This is because at the beginning of each period, when commercial loans have yet to be created, all household wealth is backed by $M_0$ and risk free bonds. Therefore, the only way for fire sales to affect real activity is via an idiosyncratic asset reallocation across households, an aggregate reallocation between liquid and illiquid asset holdings, or an aggregate reallocation of wealth between households and the government. In particular, given the zero-sum

---

68 If (1) were binding, the Walrasian constraints imposed by scarcity would lead banks to charge loan rates in excess of the otherwise prevailing competitive benchmark. In such a setting, the central bank’s task to choose appropriate amounts of high-powered money $M_0$ would become much more consequential and thus more difficult. If reserves were further required to be backed by an intrinsically valuable good (see Fama, 1980), the central bank would lose its institutional meaning entirely.
nature of nominal accounting, any loss incurred by the private sector as part of a fire sale must be offset by a corresponding profit by the central bank in my model. While the same zero-sum logic does not apply in reality, because long-term assets do in fact exist, it is nevertheless worth noting that the Federal Reserve made significant profits following its recent, large-scale asset purchases as illustrated in Figure 19.

**Figure 19.** Increased central bank profits as a symptom of unconventional monetary policy

![Graph showing increased central bank profits](image)

Notes: Figure 19 is a replication of a graph found in a Federal Reserve of St. Louis blog post (see Appendix B). It principally serves as an illustration of the ballooning central bank profits following the Federal Reserve’s extraordinary asset purchases following the 2008 financial crisis. The data were taken from FRED and FRB.

A final remark on the methodological approach

What are adequate microfoundations? The macroeconomic community has for all intents and purposes reached a consensus that its models should be microfounded, but what constitute adequate microfoundations continues to be the subject of heated debate (see Stiglitz, 2018). If macro and micro had evolved congruently, the former would be a unifying field assembling the latest behavioral insights in the form of an overarching model of the macroeconomy (Blanchard, 2018). The reason that this is not currently the case is methodological: Whenever the state space is high dimensional, there exists a tradeoff between building rich model environments and retaining canonical conditional expectation objectives. For example, with the usual disclaimer that it serves as a good point of reference, Christiano et al. (2018) poignantly proclaim that “the assumption of rational expectations is obviously wrong”. Contrarily, the proposed theory in this paper is constructed based

---

69 Since profits are returned to the treasury, government bonds purchased by the central bank effectively carry a zero percent interest rate.
on the view that the imposition of internal consistency is inadequate only if objectives are poorly chosen. In particular, since “expectations” are precisely defined mathematical objects implied by theory, instituting conditional expectation objectives requires a tremendous amount of institutional and statistical knowledge on part of the agent. Instead, I rely heavily on the literal definition of an expectation, namely “a belief that something will happen or is likely to happen” which yields the following methodological insight: There does not exist a tradeoff between insisting on internally consistent, individual optimization and building a rich macroeconomic model. In particular, I have shown the high degrees of state and parametric heterogeneity that can be accommodated when model primitives — the objectives — are chosen subject to the cognitive constraint that agents are incapable of solving Euler equations in an internally consistent manner.

But then, are the proposed microfoundations adequate? First, microfoundational adequacy does not hinge on the mathematical appeal of a model’s prescribed hypothetical behavior or the mathematical derivation thereof, but on how well the latter describes observed behavior (see Thaler and Shefrin, 1988). For example, “if Keynes was right that individuals saved a constant fraction of their income, an aggregate model based on that assumption is microfounded” (Stiglitz, 2018). However, as famously argued by Lucas (1976), even the ability to match observed behavior is insufficient to guarantee satisfactory model performance if policy is subject to change. As a result, individual behavior is now typically derived as a solution to mathematical problems of constrained optimization. Deriving behavior from optimization, rather than imposing it as a model primitive, is advantageous because it forces the theorist to disclose the fundamental economic tradeoffs that are claimed to govern the agent’s decision: “Theory helps keep track of benefits and costs” (Varian, 1993).

If the purpose of instituting objectives is to shed light on the fundamental tradeoffs considered by real-world agents, microfoundational validity not only hinges on how well prescribed behavior matches observed behavior, but also on the former’s derivation by the agents. In particular, even if a model produces decision rules that appear appealing intuitively and match the data, but no real-world agent is realistically capable to derive them, the corresponding microfoundations are unsatisfactory because the modeler for all intents and purposes imposes behavior as the model primitive and thus invariably obfuscates the actual tradeoffs considered by real-world agents. Internal consistency thus only serves as a useful benchmark if we can realistically assert that agents'
behavior indeed derives from the proposed optimization problem. In the real world, where agents are heterogenous, information is sparse, and uncertainty is epistemic, conditional-expectation objectives do not satisfy this requirement.\footnote{Neither the assumption that the infinite dimensional state is observed nor that its model implied transition is understood by real-world agents is tenable.}

In face of high degrees of heterogeneity, information sparsity, and epistemic uncertainty, there are two principal ways to address the issue of internal consistency. We can either relax the required level of rationality on part of the agents — such as boundedly rational expectation formation schemes — or, alternatively, we can reduce institutional complexity by disposing of conditional-expectation objectives. While the contemporary literature has chosen to proceed in the former fashion, I argue that reducing institutional complexity is preferable for two reasons.\footnote{See Krusell and Smith (2006) for a discussion of bounded rationality in the context of heterogenous agent setups.}

First, it preserves internal consistency as the effectively maximized objectives and the originally defined objectives are in fact congruent. Second, it improves credibility because real-world agents are almost surely incapable of deriving true conditional expectations anyway.\footnote{In ‘boundedly rational’ or ‘non-rational’ approaches, agents evaluate objectives that are incongruent with the objectives as originally defined. In this case, the proposed model primitives are ill-defined at best or agents internally inconsistent at worst. Either way, the corresponding framework inevitably violates Muth’s requirement that “expectations, since they are informed predictions of future events, are essentially the same as the predictions of the relevant economic theory” (1961).}

In this spirit, I have chosen objectives that — while still informed by the future — neither require knowledge of the entire state, nor of the full stochastic environment.\footnote{Selecting among such objectives naturally precludes the use of self-referential functional equations such as the ones that typically arise from infinite horizon setups. This is because solving the corresponding fixed point problem requires that agents, if they are to behave in an internally consistent manner, perfectly know and understand their economic and stochastic environment.} For example, rather than probabilistically assessing a stochastic sequence of future consumption, consumers assign utility to projected retirement wealth (via current liquid and illiquid savings) in my model. Instituting monetary wealth as a conceptual placeholder for future consumption makes intuitive sense because money serves, by definition, as a \textit{store of value}. Of course, the store-of-value function of money derives from the fact that future consumption is strictly increasing in accumulated monetary wealth.

Reducing institutional complexity further entails the benefit that the costs of incorporating heterogeneity are \textit{relatively} small. In particular, they are small \textit{relative} to the corresponding benefit of gaining the ability to consider cross-sectional statistics beyond the mean as macroeconomic targets (e.g. Gini coefficients). At the same time, however, allowing for heterogeneity also raises...
new questions, especially if the latter is parametric. For example, in my model, income inequality is chiefly driven by heterogenous firm ownership with skill and time-varying factors only playing secondary and tertiary roles. This may, and likely should, be viewed as problematic because there does not exist a market for firm ownership such that income inequality is for all intents and purposes determined exogenously. But what if real-world income inequality were indeed primarily driven by time-invariant firm ownership? If this were so, irrespective of whether income inequality is viewed as desirable or undesirable, incorporating heterogenous firm ownership would be key to understanding the effects of corresponding government policies.

Finally, recall that while parameterization is often viewed as successful if the resulting model generates a set of desired macro moments, credibility invariably hinges on whether the resulting parameter values are in fact consistent with microeconomic evidence (Chari, Kehoe, and McGrattan, 2009). Of course, allowing for parametric heterogeneity does not relieve the economic modeler of this credibility constraint. However, to the extent that real-world agents are in fact heterogenous in their fundamental evaluation of costs and benefits, the incorporation of parametric heterogeneity constitutes a tremendous advancement in terms of matching microeconomic evidence. Moreover, if heterogeneity is deemed contextually inappropriate, the homogenous case is, of course, still nested by the heterogenous case.

6 Concluding remarks and outlook

The key contribution of this paper lies in the formalization of a wealth effect driven aggregate demand channel in the transmission of financial crises. The proposed framework yields insights into optimal policy response while also providing a macroprudential recommendation regarding ex ante crisis prevention. In particular, since any crisis response that successfully prevents a deep downturn must do so by effectively circumventing the prevailing maintenance margin, policy makers may want to consider banning margin calls altogether. Naturally, we would expect that such a policy would increase the cost of leverage, which may or may not be viewed as desirable. Since the proposed framework does not generate any financial crises beyond liquidity induced fire sales, the logic here only applies to the prevention of crises of this type. To allow for high degrees of state and parametric heterogeneity, I have further proposed a non-standard methodological approach
that retains internal consistency and global solutions. Specifically, in an attempt to enhance the credibility of the employed microfoundations, each agent’s objective was chosen subject to the constraint that optimization must be trivial to achieve numerically.

In terms of future research, the proposed methodological approach opens up a natural avenue for analyzing cross-sections dynamically, a capability which the present paper has hardly even scratched the surface of. Going forward, I thus envision endowing households with a vector of sector-specific productivity and allowing firms to observe certain worker characteristics as in Weiss (1980). In addition to the taste shock, demand uncertainty could arise from a spatial reallocation of consumers along a circular city after production. Firms may further possess heterogeneous, potentially endogenous production technologies. Relating to technology, since modern likelihood methods are not subject to the traditional ergodicity constraints imposed by the method of moments, I further ultimately envision a growing economy that unifies trend and cycles akin to King, Plosser, and Rebelo (2002). Incorporating non-stationarity is intriguing because estimation would neither require nor allow for the atheoretical extraction of “trend” from the data.
A Equilibrium

A.0 Definition

Since the economy analyzed herein features market clearing failures in equilibrium, I start by outlining the set of relevant definitions. For this, consider a pseudo-game, or an abstract economy, in which “an action by one agent affects both the payoff and the domain of actions of other agents” (Arrow and Debreu, 1954). More formally, an abstract economy \( \Gamma_A \) emerges if some game \( \Gamma = [I, \{\Sigma_i\}_{i \in I}, \{u_i\}_{i \in I}] \) is augmented by a set of correspondences \( \{A_i\}_{i \in I} : \Sigma_{-i} \rightrightarrows \Sigma_i \).

**Pseudo-Nash equilibrium.** A strategy profile \( \sigma = \{\sigma_i\}_{i \in I} \) is said to be a pseudo-Nash equilibrium (PNE) of an abstract economy \( \Gamma_A = [I, \{\Sigma_i\}_{i \in I}, \{u_i\}_{i \in I}, \{A_i\}_{i \in I}] \) if \( u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma'_i, \sigma_{-i}) \) for all \( i \in I, \sigma'_i \in A_i(\sigma_{-i}) \).

**Competitive equilibrium.** A specific market within an abstract economy is said to be competitive if pseudo-Nash equilibrium induces zero excess demand in said market. An abstract economy is said to be competitive if pseudo-Nash equilibrium induces zero excess demand in each market.

**Walrasian equilibrium.** A competitive market is said to be Walrasian if there exists one market participant who’s sole purpose is to act as the market’s Walrasian auctioneer. An abstract economy is said to be Walrasian if each market features a Walrasian auctioneer.

Crucially, notice that pseudo-Nash equilibrium may not be competitive such that the prevalence of market clearing failures need not imply disequilibrium in the pseudo-Nash sense. In fact, since not all markets in the recursive economy examined herein are competitive, the relevant equilibrium concept, recursive general equilibrium, must explicitly permit excess demand, positive or negative.

**General equilibrium.** An abstract economy is said to be of the general equilibrium type if the price in each market is determined endogenously by pseudo-Nash equilibrium.

**Recursive general equilibrium.** A recursive model economy is said to be of the general equilibrium type (RGE) if each price in each market is determined endogenously by pseudo-Nash equilibrium in (each subperiod of) each period.\(^{75}\)

---

\(^{74}\)e.g. the price set by a Walrasian auctioneer restricts the set of permitted strategies by the market participants.

\(^{75}\)Notice that RGE encompasses both “dynamic stochastic” (DSGE) as well as “recursive-dynamic computable” (RDCGE) setups. It is well known that the combination of infinite horizons with rational expectations, a defining feature of DSGE, requires the model to be solved across a grid of possible states, a computational feat that is subject to
In addition to market clearing failures, notice that the proposed model economy features markets that are highly incomplete. Specifically, this is because each agent is restricted to engage in specific markets corresponding to her type. Moreover, trading exclusively takes place in recurrent spot markets and real goods can only be acquired in exchange for $M_1$.

A.1 Market for consumption goods (and aggregate demand)

Once production has occurred, all production costs are sunk. At $t_4$, the observation of the taste shock $\xi_t$ pins down aggregate demand $D(P)$ and thus each firm’s residual demand as a function of the vector $\{P^C_{jt}, y^C_{jt}\}_{j \neq f}$. It is assumed that, at this stage, firms compete in Bertrand fashion and that, if multiple firms offer the same price, residual demand is allocated proportionally. We have,

$$\max_{P^C_{jt}} P^C_{ft} \min \left\{ y^C_{ft}, y^C_{d,ft} \right\}$$

where residual demand is given by $y^C_{d,ft} = D(P^C_{ft}) - \sum_{j \neq f} y^C_{jt} \mathbb{1}(P^C_{jt} < P^C_{ft}) - \sum_{j \neq f} \frac{y^C_{jt} \mathbb{1}(P^C_{jt} = P^C_{ft})}{y^C_{jt} + y^C_{ft}}$. First, notice that no firm will optimally choose to set their price below the market clearing price $P^* = \{P \in \mathbb{R} \mid D(P) = \sum_{j \neq f} y^C_{jt}\}$ at which point $y^C_{ft}$ is sold in its entirety irrespective of the competition’s pricing. However, depending on other firms’ offering prices, a firm may find it profitable to charge a price above $P^*$, $P^C_{ft} = D^{-1}(y^C_{ft})$ if $P^C_{jt} > D^{-1}(y^C_{ft}) \ \forall j \neq f$ for example.

Before proceeding to discuss optimal pricing, I show that sales $D(P)P$, or household expenditures, are weakly decreasing in price. For this, suppose the lowest available price $P^C_0 = \min_{f}\{P_{ft}\}$ induces a particular household to save $s_0$ and consume $c_0 = w^L(1 - s_0)/P^C_0$. Assuming an interior solution, we must then have,

---

76 No agent trades firm ownership or old capital because no such markets exist.

77 See Kreps and Scheinkman (1983) for a more general treatment.
\[
\gamma c_0 \gamma^{-1} \frac{w^L}{P_0} = \xi_t \exp(-\gamma w^{G}(s_0)) a \\
\equiv m(s_0)
\]

where \( a > 0 \) is a constant such that \( m' < 0 \) and, crucially, \( s_0 \) is the only non-predetermined argument of \( m \). Now, suppose that the price increased to \( P_1^C = \lambda P_0^C \), \( \lambda > 1 \), but that the household responds by saving \( s_1 \leq s_0 \). We then have \( c_1 \geq c_0/\lambda \), which in turn implies,

\[
\gamma c_1 \gamma^{-1} \frac{w^L}{P_1^C} \leq \gamma \left( \frac{c_0}{\lambda} \right) \gamma^{-1} \frac{w^L}{\lambda P_0^C} \\
= \lambda^{-\gamma} \gamma c_0 \gamma^{-1} \frac{w^L}{P_0^C} \\
= \lambda^{-\gamma} m(s_0) \\
< m(s_0) \\
\leq m(s_1)
\]

where the last inequality follows by \( m' < 0 \). Therefore, \( P_1^C > P_0^C \) implies that the marginal benefit of saving strictly exceeds its marginal cost at any \( s_1 \leq s_0 \). Unless the respective household was already at the corner \( s_0 = 0 \), it will thus respond by increasing its savings or, equivalently, decreasing its expenditures such that we must have \( s_1 > s_0 \). Since the same logic applies equally across all households, we have \( \partial D(P^C) P^C / \partial P^C \leq 0 \) over the entire domain of \( P^C \). Moreover, on the subset of the domain where there exists at least one household not at the corner \( s_0 = 0 \), which must be satisfied for at least one realization of the taste shock in equilibrium\(^78\), the inequality is strict: \( \partial D(P^C) P^C / \partial P^C \big|_{P^*} < 0 \) where \( P^* \) denotes the market clearing price in general equilibrium. Incorporating this insight in their estimation of market demand, firms rightfully impose \( \tilde{\chi}_t^r(\xi_t) \leq 1 \) for each \( \xi_t \). I now proceed to show that equilibrium at \( t_4 \) is given by a uniform market clearing price strategy,

\[ P_{ft}^C = P^* \text{ for each } f \in J^C \]

By contradiction: Suppose that, in equilibrium, \( \exists f \) such that \( P_{ft}^C > P^* \). Then, if \( y_{ft}^C < y_{ft}^{C,d} \),

\(^78\)Otherwise, firms trivially have an incentive to produce less.
firm $f$ reacts by raising its price. If $y_{ft}^C = y_{ft}^{Cd}$, some firm $j$ has zero sales and thus optimally reacts by lowering its price (at least to $P^C_{ft} - \varepsilon$, maybe further). If $y_{ft}^C > y_{ft}^{Cd} > 0$, firm $f$ is a marginal seller and thus has an incentive to lower prices because its projected sales are at least locally decreasing in price (sales increases are discrete if there are other marginal sellers). Similarly, if $y_{ft}^C > y_{ft}^{Cd} = 0$, firm $f$ will lower its price to attain positive sales.

### A.2 Production

As indicated, firms do not understand that they can affect other firms’ output by poaching workers. Taking as given the announced firm outputs, optimality requires,

$$
\left( \frac{\partial}{\partial k_{ft}^C} \frac{\partial}{\partial n_{ft}^C} \right) \left( E_{t1}^C \left[ \tilde{S}_{ft}^C \right] - \left[ k_{ft}^C Q_t + \mu^L F n_{ft}^C W_{ft}^C R_{ft}^L \right] \right) = (0, 0)
$$

which implies

$$
\frac{\partial E_{t1}^C \left[ \tilde{S}_{ft}^C \right]}{\partial y_{ft}^C} \frac{\partial y_{ft}^C}{\partial k_{ft}^C} / \frac{\partial E_{t1}^C \left[ \tilde{S}_{ft}^C \right]}{\partial y_{ft}^C} \frac{\partial y_{ft}^C}{\partial n_{ft}^C} = \frac{Q_t R_{ft}^L}{\mu^L F W_{ft}^C R_{ft}^L}
$$

$$
\implies \frac{\partial y_{ft}^C}{\partial k_{ft}^C} / \frac{\partial y_{ft}^C}{\partial n_{ft}^C} = \frac{Q_t}{\mu^L F W_{ft}^C}
$$

and therefore, given the Cobb-Douglas form of production, constant expenditure shares for capital and labor,

$$
k_{ft}^C Q_t = \frac{\alpha}{1 - \alpha} \mu^L F n_{ft}^C W_{ft}^C
$$

(11)

Since production is CRS and prices are taken as given, the marginal cost of producing additional units of output is equal to a constant $\delta_t^C$. To find $\delta_t^C$, I exploit the optimal capital-labor share and calculate the cost of producing a benchmark output with $n_{ft}^C = 1$,

$$
\delta_t^C = \left( \frac{\mu^L F q_{ft}^C}{(1 - \alpha) Q_t} \right) Q_t + \mu^L F W_{ft}^C \right) R_{ft}^L
$$

$$
= \frac{Q_t^\alpha [W_{ft}^C]^{1-\alpha} R_{ft}^L}{\alpha^\alpha (1 - \alpha)^{1-\alpha} \tilde{S}_{ft}^C [q_{ft}^C]^{1-\alpha}}
$$

such that the marginal cost is linear in $R_{ft}^L$.\(^{79}\) Consumption sector firms thus implicitly optimize,

\(^{79}\)As per usual, monetary policy does not have any real effects unless nominal frictions prevent markets from
max \( \mathbb{E}_t \left[ y_{ft}^C \bar{X}_t^k(\xi_t) \left[ \sum_{jC} y_{jt}^C \bar{X}_t^r(\xi_t) \right] - \delta_t^C y_{ft}^C \right] \)

Appealing to symmetry, it is then easy to show that in equilibrium, individual output solves the following equation,

\[
\mathbb{E}_t \left[ \tilde{X}_t^k(\xi_t) y_{ft}^C - \tilde{X}_t^r(\xi_t) \left( 1 - \frac{\tilde{X}_t^r(\xi_t)}{N^C} \right) \right] = \delta_t^C
\]

such that equilibrium output is, ceteris paribus, increasing in \( \tilde{X}_t^k \) and decreasing in \( \tilde{X}_t^r \). Further notice that (10) can be written as,

\[
\mathbb{E}_t \left[ \tilde{p}_t^C(\xi_t) \left( 1 - \frac{\tilde{X}_t^r(\xi_t)}{N^C} \right) \right] = \delta_t^C
\]

which reduces to the familiar zero expected profits condition as \( N^C \to \infty \).

A.3 Market for capital

I start by showing that consumption good sector equilibrium dictates that consumption producer capital expenditures \( k_t^C(Q_t)Q_t \), and thus capital producer rental revenue, must be strictly decreasing in \( Q_t \). From the firm’s FOC for capital, we know,

\[
\alpha \mathbb{E}_t \left[ \tilde{X}_t^k(\xi_t) [N^C y_{ft}^C]^{-\tilde{X}_t^r(\xi_t)} \left( 1 - \frac{\tilde{X}_t^r(\xi_t)}{N^C} \right) \right] y_{ft}^C = k_t^C Q_t R_{ft}^L
\]

In comparing different equilibria (taking as given two different values of \( Q_t \)), it is useful to analyze the behavior of both sides of (10). Specifically, notice that \( \tilde{X}_t^k(\xi_t) \leq 1 \) implies that the left hand side of (10) is (at least weakly) increasing in \( y_{ft}^C \). Similarly, the left hand sides of (11) and (12) are strictly decreasing in \( k_t^C \) and \( n_t^C \) respectively,

\[
\alpha \mathbb{E}_t \left[ \tilde{X}_t^k(\xi_t) [N^C y_{ft}^C]^{-\tilde{X}_t^r(\xi_t)} \left( 1 - \frac{\tilde{X}_t^r(\xi_t)}{N^C} \right) \right] \frac{y_{ft}^C}{k_t^C} = Q_t R_{ft}^L
\]

\[
(1 - \alpha) \mathbb{E}_t \left[ \tilde{X}_t^k(\xi_t) [N^C y_{ft}^C]^{-\tilde{X}_t^r(\xi_t)} \left( 1 - \frac{\tilde{X}_t^r(\xi_t)}{N^C} \right) \right] \frac{y_{ft}^C}{n_t^C} = \mu^{LF} W_{ft}^C R_{ft}^L
\]

clearing. In particular, when prices are entirely flexible, all exogenous changes in \( R_t^R \) are absorbed by a corresponding joint level shift in \( W_t^C \) and \( Q_t \) in equilibrium. However, since fire sales are transmitted to the real sector via the nominal downward wage friction, the central bank can mitigate the real effects of a crisis by lowering the interest rate target and thereby absorbing part of the nominal shock. Conversely, the monetary authority could hypothetically also generate unemployment by sharply increasing the interest rate target at any given time.
By contradiction: Suppose $Q_0$ and $Q_1 > Q_0$ give rise to two firm-bank equilibria as given by the individual strategies $(W_0, n_0, k_0, l_0, y_0; R_0)$ and $(W_1, n_1, k_1, l_1, y_1; R_1)$ satisfying $k_1Q_1 \geq k_0Q_0$. From (8), we then know that $n_1W_1 \geq n_0W_0$ and thus $l_1 \geq l_0$. Without loss of generality, assume then that $l_1 = a_0$ with $a \geq 1$. Given $Q_1 > Q_0$, we must have $k_1 < ak_0$. Moreover, from the labor market setup we know that, in equilibrium, $n_1 > n_0 \implies W_1 \geq W_0$ such that, because $n_1W_1 = an_0W_0$, we must have $n_1 \leq an_0$. Combining $k_1 < ak_0$ and $n_1 \leq an_0$ yields $y_1 < ay_0$, which implies, by $\delta^C_t > 0$, that expected (projected) sales per unit of the loan have decreased, $\mathbb{E}_{t_t} [S(y_1)]/l_1 < \mathbb{E}_{t_t} [S(y_0)]/l_0$. Then, unless there exists no $\xi_t \in \Xi$ such that sales fall short of $l_0R_0$\(^8\), whether $y_1 < y_0$ or $y_1 \leq y_0$, we must have $R_1 > R_0$. However, notice that $y_1 > y_0$ implies $k_1 > k_0$ or $n_1 > n_0$ and thus, by (11) or (12), $R_1 < R_0$. Similarly, $y_1 \leq y_0$ implies, by (10), $R_1 \leq R_0$. We have a contradiction. Therefore, since optimality requires $\partial k_t^C (Q_t) Q_t / \partial Q_t < 0$ in equilibrium at the individual firm level, we have $\partial k_t^C (Q_t) Q_t / Q_t < 0$ in aggregate.

I proceed by showing that equilibrium in the capital goods sector is characterized by a uniform, market clearing capital rental price $Q_{ft} = Q^*_t$ for each $f$ and a sectoral wage schedule of $W^K_{ft} (y^K_{ht}) \equiv \max_f \left\{ W^K_{ft} (y^K_{ht}) \right\} = y^K_{ht}Q_t$ for each $h$. Recall that the capital goods producing firms maximize contemporaneous monetary profits by choosing an individual price $Q_{ft}$ and an individual wage offer $W^K_{ft} (y^K_{ht})$.

\[
V^K_{ft-1} (k^K_{ft-1}, \{k^K_{jt-1}, P^K_{jt}, W^K_{jt} \}_{j \neq f}) = \max_{Q_{ft}, W^K_{ft}} Q_{ft} \min \left\{ k^K_{ft}^d, k^K_{ft}^s \right\} - \int n^K_{ht} W^K_{ft} (y^K_{ht}) \, dh
\]

subject to $k^K_{ft} = (1 - \delta^D) k^K_{ft-1} + y^K_{ft}$

\[
y^K_{ft} = \int n^K_{ht} y^K_{ht} \, dh
\]

where $k^K_{ft}^d = k^K_t (Q_{ft}) - \sum_j k^K_{jt} 1 (Q_{jt} < Q_{ft}) - \sum_j c \frac{k^K_{jt} (Q_{jt} = Q_{ft})}{k^K_{jt} + k^K_{jt}^s}$. I start by discussing existence constructively: suppose that for each $h$, the highest available wage contract, offered by at least two firms, is given by $W^K_{ft} (y^K_{ht}) = y^K_{ht}Q^*_t$, where $Q_{ft} = Q^*_t$ for each $f$ and $Q^*_t$ clears the market. Then, each employed worker generates zero marginal profits and has a competitive outside option. Therefore, lowering wage offers cannot be profitable because current employees will simply opt to work for another firm. Conversely, poaching a worker from another firm by offering a higher

---

\(^8\)If there were no risk associated with $l_0$ and $l_t$, we would have $R_0 = R_1 = R^S$, but this is contextually irrelevant as the primary purpose of the taste shock is to create such risk.
wage may increase a firm’s output, but only at the cost of negative marginal profits: If $Q_{ft}$ is left unchanged (or lowered), all product is still sold, but the marginal sale does not cover the marginal labor costs. If $Q_{ft}$ is raised above $Q^*_t$, markets no longer clear. Increased total labor costs are then mirrored by shrinking sales because, as shown above, capital rentals $k^C_t(Q_t)Q_t$ are strictly decreasing in $Q_t$ if the consumption goods sector is in equilibrium.

Having shown existence, I proceed by proving uniqueness. For this, notice that, in equilibrium, all capital rentals must occur at a uniform price. More precisely, $Q_{ft} = Q_t$ for all $f$ satisfying $k^S_{ft} > 0, k^D_{ft}(Q_{ft}) > 0$. By contradiction: Suppose that in equilibrium, $\exists i,j$ such that $Q_{it} > Q_{jt}$ and $k^S_{it}, k^D_{jt}(Q_{it}) > 0$. Then, if markets clear, $j$ may, irrespective of wages paid, increase profits by raising $Q_{jt}$ to $Q_{it}$. In case of excess demand, both firms find it profitable to raise their price to the market clearing price $Q^*_t$. Finally, in case of excess supply, $j$ may increase profits by raising $Q_{jt}$ to $Q_{it} - \varepsilon$. Since existence has been shown and price dispersion cannot support equilibrium, all capital rentals must occur at a uniform price $Q_t$ in equilibrium. Given $Q_t$, a worker $h$ must be offered $W^K_t(y^K_{ht}) = y^K_{ht}Q_t$ in equilibrium. By contradiction: Suppose that, in equilibrium, $\exists h$ such that $W^K_t(y^K_{ht}) = \tilde{W}^K_t(y^K_{ht}) \neq y^K_{ht}Q_t$. If $W^K_t(y^K_{ht}) < y^K_{ht}Q_t$, at least one firm, to attract the profitable worker $h$, will deviate by offering $W^K_{ft}(y^K_{ht}) = W^K_t(y^K_{ht}) + \varepsilon$. Conversely, if $W^K_t(y^K_{ht}) > y^K_{ht}Q_t$, the hiring firm would incur a net loss by employing $h$ such that it is more profitable to lower its offer until $W^K_{ft}(y^K_{ht}) \leq y^K_{ht}Q_t$. Given $Q_{ft} = Q_t$ for all $f$ satisfying $k^S_{ft} > 0, k^D_{ft}(Q_{ft}) > 0$ and $W^K_{ft}(y^K_{ht}) = y^K_{ht}Q_t$, suppose that $Q_t$ does not clear the market. Then, in case of excess demand, it is trivially profitable for each firm to raise prices to $Q^*_t$. In case of excess supply, since equilibrium capital expenditures $k^C_t(Q_t)Q_t$ are strictly decreasing in $Q_t$, firms find it profitable to lower their price until the market clears at $Q^*_t$.

A.4 Market for labor

Workers supply labor to the consumption goods sector if the unemployment benefits exceed their respective outside option $\delta^U W^C_t > q_hz^K_tQ_t$. Aggregate labor supply is then given by,

$$n^C_s = \Pr (q_h < \delta^U W^C_t / z^K_tQ_t)$$

$$= G \left( \delta^U W^C_t / z^K_tQ_t \right)$$

In other words, the wage necessary to attract any given worker is, ceteris paribus, increasing in
On the other side of the market, labor demand satisfies,

\[ n^C_t = \sum_{j^C} n^C_{jt} \]

\[ = \sum_{j^C} \frac{(1 - \alpha)k^C_{jt}Q_t}{\mu LF \alpha W^C_t} \]

\[ = \frac{(1 - \alpha)k^C_{t}Q_t}{\mu LF \alpha W^C_t} \]

Since equilibrium requires \( W^C_t \geq \max \{ W^C, \delta W^C_{t-1} \} \), the labor market may exhibit excess supply in equilibrium. For example, consider Figure 20 which depicts a situation where the labor market fails to clear.

**Figure 20.** Retirement savings guidance from the US Department of Labor

The market for labor may exhibit excess supply \( n^C_{s,t} > n^C_t \) for two reasons. First, firms never find it profitable to lower wages by more than \( 1 - \lambda^W \) relative to last period’s wage \( W^C_{t-1} \) because of concerns relating to worker effort. Second, the fact that offering higher wages increases average labor productivity induces the lower wage threshold \( W^C_t \). \(^{82}\)

Following the procedure of Weiss (1980), suppose there exists a wage offer \( W^o < \frac{W^C}{z_t Q_t} \), a labor demand \( n^o \), and a corresponding output \( y^o \) that maximize the firm’s payoff. Holding

\(^{81}\)Since the equilibrium wage rate \( W^C_t \) is decreasing in \( \delta^U \), both \( n^C_{s,t} \) and \( n^C_t \) are increasing in \( \delta^U \) in equilibrium. Therefore, if the demand shift dominates the supply shift, raising unemployment benefits induces lower equilibrium unemployment.

\(^{82}\)Technically, the described adverse selection mechanism is not enough to generate the lower bound. To generate the lower bound, I require that (2) has an interior maximum, which requires that average productivity also be locally concave. The desired concavity property is provided by \( G_q \) being rectified Gaussian, a rather unrealistic assumption. For future work, I envision introducing scale efficiencies via nonlinear skill aggregation.

62
labor costs $C^o = W^o n^o$ fixed, the firm may alternatively employ $n^* = C^o / W^C < n^o$ workers at $W^C$. From (2), we know that $q^C(W^C)/W^C > q^C(W^0)/W^0$ and thus, since labor costs are fixed, $q^C(W^C)n^* > q^C(W^o)n^o$ in which case the firm produces $y^* > y^o$ at the same cost $C^o$. By continuity of production and the strictly decreasing labor cost, the firm may alternatively also produce $y^c$ at a reduced labor cost. Therefore, as long as the firm maximizes some measure of contemporaneous profit, we have induced a contradiction and thus shown that offering any $W^0 < W^C$ is strictly dominated by the strategy of offering $W^C$ and employing $n^*$. 

A.5 Market for money (demand deposits and reserves)

Banks are special in my model because they can artificially create previously non-existing units of the economy’s numéraire by lending to firms. Since they are risk averse, banks never opt to keep a commercial loan on balance sheet, but they have the ability to infer what price $P^L_{bft}$ can be obtained from selling a particular commercial loan $(l_{bft}, R^L_{bft})$ on the secondary market. Because competition must yield zero profits, we have $l_{bft} = P^L_{bft}$. Following creation, the newly minted demand deposits vanish as soon as the SPV contracts are executed. Finally, notice that banks also engage in the market for Fed Funds to satisfy the prevailing reserve requirement with the outside option of investing in the risk free bond pinning down $R^S_{FFR} = R^S_t$. 

A.6 Market for debt and equity financing

Given the auctioneer’s announcement $R^D_t$, households maximize a risk adjusted measure of projected asset returns, whereas funds maximize projected return on equity. The household’s optimality condition is given by,

$$\tilde{R}^E_t - \tilde{R}^D_t \geq \gamma_h \frac{w^E_{hta}}{w^I_{hta}}$$

which implies,

$$\frac{w^E_{hta}}{w^I_{hta}} = \min \left\{ \frac{1}{\gamma_h (\tilde{R}^E_t - \tilde{R}^D_t)}, 1 \right\}$$

such that the optimal equity share is increasing in the projected risk premium $\tilde{R}^E_t - \tilde{R}^D_t$ and decreasing in $\gamma_h$. Finally, the equilibrium interest rate $R^D_t$ clears the market,

$$\int_{J_{HH}} w^E_{hta} = \frac{1}{1 + \delta_t} \int_{J_{HH}} w^I_{hta}$$
B Data

The model was parameterized using quarterly US data from 1987 until 2017. The analysis is limited to this time period because the institutional monetary policy changes undertaken by former Fed chair Volcker are widely believed to have muted the business cycle (see Stock and Watson, 2002). Therefore, since the inclusion of structural changes invariably undermines the ubiquitous proposition of ergodicity, moment matching is conducted using data from the post-Volcker era, also known as the Great Moderation, only. All labor market data was obtained from the US Bureau of Labor Statistics, whereas the other series were sourced from the Federal Reserve of St. Louis database (FRED), the Federal Reserve Board (FRB), or Yahoo Finance. The following contains the origin of all data series displayed in the respective figures.

**TABLE 7. DATA†**

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Series</th>
<th>Frequency (level), source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Unemployment</td>
<td>Monthly (2008=0), LNS14000000 via BLS</td>
</tr>
<tr>
<td></td>
<td>Employment</td>
<td>Monthly (2008=0), LNS12300000 via BLS</td>
</tr>
<tr>
<td></td>
<td>Employment, 25-54</td>
<td>Monthly (2008=0), LNS12300060 via BLS</td>
</tr>
<tr>
<td>2.</td>
<td>Job losers, on layoff</td>
<td>Monthly, LNS13023653 via BLS</td>
</tr>
<tr>
<td></td>
<td>Job losers, others</td>
<td>Monthly, LNS13025699 via BLS</td>
</tr>
<tr>
<td></td>
<td>Job leavers (quit)</td>
<td>Monthly, LNS13023705 via BLS</td>
</tr>
<tr>
<td></td>
<td>Reentrants</td>
<td>Monthly, LNS13023557 via BLS</td>
</tr>
<tr>
<td></td>
<td>New entrants</td>
<td>Monthly, LNS13023569 via BLS</td>
</tr>
<tr>
<td>3.</td>
<td>Unemployment</td>
<td>Quarterly, LNS14000000 via BLS</td>
</tr>
<tr>
<td>5.</td>
<td>Answers</td>
<td>Quarterly from Amir Sufi’s Website</td>
</tr>
<tr>
<td>6.</td>
<td>Employment</td>
<td>Monthly, LNS12300060 via FRED</td>
</tr>
<tr>
<td></td>
<td>Wages</td>
<td>Monthly, AHETPI via FRED</td>
</tr>
<tr>
<td></td>
<td>CPI</td>
<td>Monthly, PCEPILFE via FRED</td>
</tr>
<tr>
<td>13.</td>
<td>FFR</td>
<td>Monthly, DFF via FRED</td>
</tr>
<tr>
<td></td>
<td>1Y Treasury</td>
<td>Monthly, GS1 via FRED</td>
</tr>
<tr>
<td>19.</td>
<td>Transfers</td>
<td>Annual, B1190C1A027NBEA via FRED and FRB</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>Annual, GDPA via FRED</td>
</tr>
<tr>
<td></td>
<td>Federal receipts</td>
<td>Annual, FYFR via FRED</td>
</tr>
</tbody>
</table>

†Notes: Figure 3 plots a Gaussian kernel density estimate of US unemployment since 1987. The estimate is constructed as in Botev et al. (2010) with a mesh granularity of $2^{-7}$. The employed input frequency is reduced to quarterly because the model is parameterized to match quarterly data. Figure 19 is a replication of a graph found in the Federal Reserve of St. Louis blog post “Fed Payments to Treasury and Rising Interest Rates” by Miguel Faria-e-Castro.
C Top 10 Ways to Prepare for Retirement

In my framework, households do not maximize expected lifetime utility over an infinite stream of consumption. First, this is because they only live a finite number of periods. More importantly, it’s because deriving the distribution over future consumption implied by the model is extraordinarily challenging from a computational perspective. However, households understand that future consumption, retirement consumption in particular, is strictly increasing in accumulated lifetime savings. Rather than worrying about future consumption directly, my objective’s relevant marginal benefit associated with the marginal cost of decreasing consumption today is thus given by the projected increase in accumulated retirement balances. This modeling choice entails the cost that future consumption only generates utility implicitly — because retirement consumption is increasing in retirement balances — such that the resulting decision rule is only implicitly governed by the tradeoff between consumption today and consumption in the future. Moreover, each household’s future plan is reduced to the assumption that they will simply continue to save the exact same nominal amount each period until retirement. Admittedly, the proposed objective constitutes a stark departure from contemporary practice, but it is motivated by the real-life mathematical exercise depicted in Figure 21, which is frequently used, as illustrated by the fact that the US Department of Labor advertises it, to illustrate the benefits associated with saving for retirement early.\textsuperscript{83}

Figure 21. Retirement savings guidance from the US Department of Labor

Notes: The above illustration is taken from the US Department of Labor’s publication “Top 10 Ways to Prepare

\textsuperscript{83}In essence, the exercise illustrates the (nominal) marginal benefit of saving more today, again assuming that one saves the exact same amount each period until retirement.
for Retirement”. It serves as an inspiration as to how households effectively assess the canonical tradeoff between consuming today and consuming in the future. In particular, I assume that households find probabilistic assessments of future consumption too complex and thus resort to setting themselves a nominal retirement savings goal. To achieve this goal, they go through the above cumulative compounding exercise while appealing to an ergodic benchmark interest rate.

As is the case in reality, pinning down a reasonable benchmark for the relevant interest rate used for compounding presents a challenging task. In the model, I assume that households view monetary policy shocks as permanent, namely by interpreting the risk free rate as ‘the current interest rate level’, and derive current asset return projections by adding the (ergodic) mean of all historic markups to the prevailing risk free rate.

D SMM and ergodicity

Consistency of SMM requires that the proclaimed data generating process $X : \Omega \times T \mapsto S$, defined on $(\Omega, W, \mu_W)$, whose moments we want to match is ergodic.\(^{84}\)

**Geometric ergodicity.** A process $X : \Omega \times T \mapsto S$ is said to be geometrically ergodic if there exists $\rho < 1$ and a time-invariant measure $f_X : S \mapsto \mathbb{R}$ such that for any $x_0 \in S$,

$$
\lim_{T \to \infty} \rho^{-T} \|P_T^{x_0} - f_X\|_\nu = 0
$$

where $\|\cdot\|_\nu$ is the total variation norm and $P_T^{x_0} : S \mapsto \mathbb{R}$ denotes the conditional distribution of $X_T$ given $X_0 = x_0$. Assuming geometric ergodicity and some regularity conditions (see Duffie and Singleton, 1993), we have,

$$
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} g(X_t) = \lim_{T \to \infty} \mathbb{E}[g(X_T)] = \int_S g(x) df_X(x)
$$

for any $f_X$-measureable function $g$ and any initial condition $X_0 = x_0 \in S$. In particular, notice that aside from targeting canonical first and second moments — via the identity map and the squared error map — $g$ may be chosen target any ergodic statistic including conditional expectations and quantiles. Letting $X_\infty$ denote any random variable with the density $f_X$, consider for example $g(x) = g_c(x) \equiv x 1(x \leq q)$,

\(^{84}\)In calculating asymptotic distributions, (geometric) ergodicity can substitute for stationarity since it means that the process converges (geometrically) to its stationary distribution” (Duffie and Singleton, 1993).
\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} X_t \mathbb{I}(X_t \leq q) = \lim_{T \to \infty} \mathbb{E}[X_T | X_T \leq q] \\
= \mathbb{E}[X_\infty | X_\infty \leq q]
\]

where \( \bar{T} \equiv \sum_{t=0}^{T} \mathbb{I}(X_t \leq q) \). Alternatively, consider \( g(x) = g_q(x) \equiv \mathbb{I}(x \leq q) \),

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}(X_t \leq q) = \lim_{T \to \infty} \mathbb{E} [\mathbb{I}(X_T \leq q)] \\
= \lim_{T \to \infty} \mathbb{P}(X_T \leq q) \\
= \mathbb{P}(X_\infty \leq q) \\
= F_{X_\infty}(q)
\]

such that the LHS provides a consistent estimate of the \( q \)th quantile of the limiting cdf \( F_X \).

E Institutional dictionary

<table>
<thead>
<tr>
<th>TABLE 8. AGENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meaning</strong></td>
</tr>
<tr>
<td><strong>Banks</strong></td>
</tr>
<tr>
<td><strong>Brokers</strong></td>
</tr>
<tr>
<td><strong>Capital producers</strong></td>
</tr>
<tr>
<td><strong>Consumption goods producers</strong></td>
</tr>
<tr>
<td><strong>Pension funds</strong></td>
</tr>
<tr>
<td><strong>Households</strong></td>
</tr>
<tr>
<td><strong>Meaning</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Calibration</td>
</tr>
<tr>
<td>Commercial loan</td>
</tr>
<tr>
<td>Expectation</td>
</tr>
<tr>
<td>Estimation</td>
</tr>
<tr>
<td>Ergodicity</td>
</tr>
<tr>
<td>Fire Sale</td>
</tr>
<tr>
<td>Fragility</td>
</tr>
<tr>
<td>Initial margin</td>
</tr>
<tr>
<td>Maintenance margin</td>
</tr>
<tr>
<td>Margin call</td>
</tr>
<tr>
<td>Monetary policy, conventional</td>
</tr>
<tr>
<td>Monetary policy, unconventional</td>
</tr>
<tr>
<td>Projection</td>
</tr>
<tr>
<td>Pricing, absolute</td>
</tr>
<tr>
<td>Pricing, relative</td>
</tr>
<tr>
<td>Sellout</td>
</tr>
<tr>
<td>Macroeconomic transmission</td>
</tr>
</tbody>
</table>
References


71


