NBER WORKING PAPER SERIES

IMPERFECT RISK-SHARING AND THE BUSINESS CYCLE

David W. Berger
Luigi Bocola
Alessandro Dovis

Working Paper 26032
http://www.nber.org/papers/w26032

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2019, Revised February 2023

First draft: February 2019. We thank Fernando Alvarez, Adrien Auclert, Anmol Bhandari, Loukas Karabarbounis, Greg Kaplan, Patrick Kehoe, Dirk Krueger, Guido Lorenzoni, Alisdair McKay, Kurt Mitman, Ben Moll, Luigi Pistaferri, Richard Rogerson, Joe Vavra and participants at several seminars and conferences. We thank Aniket Baksy and Sara Casella for excellent research assistance. Luigi Bocola thanks the Alfred P. Sloan foundation for generous funding. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Imperfect Risk-Sharing and the Business Cycle
David W. Berger, Luigi Bocola, and Alessandro Dovis
NBER Working Paper No. 26032
July 2019, Revised February 2023
JEL No. E32,E44

ABSTRACT

This paper studies the macroeconomic implications of imperfect risk sharing implied by a class of New Keynesian models with heterogeneous agents. The models in this class can be equivalently represented as a representative-agent economy with wedges. These wedges are functions of households' consumption shares and relative wages, and they identify the key cross-sectional moments that govern the impact of households' heterogeneity on aggregate variables. We measure the wedges using U.S. household-level data, and combine them with a representative-agent economy to perform counterfactuals. We find that deviations from perfect risk sharing implied by this class of models account for only 7% of output volatility on average, but can have sizable output effects when nominal interest rates reach their lower bound.

David W. Berger
Department of Economics
Duke University
419 Chapel Drive
Social Science Building 231
Durham, NC 27708
and NBER
david.berger@duke.edu

Luigi Bocola
Department of Economics
Landau 342
Stanford University
579 Jane Stanford Way
Stanford, CA 94305-6072
and NBER
lbocola@stanford.edu

Alessandro Dovis
Department of Economics
University of Pennsylvania
The Ronald O. Perelman Center
for Political Science and Economics
133 South 36th Street
Philadelphia, PA 19104
and NBER
adovis@upenn.edu

1 Introduction

To what extent are households' heterogeneity and deviations from perfect risk sharing important for aggregate fluctuations? Building on the influential quasi-aggregation result in Krusell and Smith (1998), for long time the consensus view in macroeconomics was that these features were not critical drivers of the business cycle. With the Great Recession, however, several economists suggested that shocks and frictions at the households' level, such as tighter credit constraints or heightened income risk, could foster households' precautionary savings and explain the observed persistent slump in aggregate demand. Despite recent advances in understanding how these economic forces play out in general equilibrium models, quantifying their role is still an open question. This is because the impact that these microeconomic frictions have on macroeconomic aggregates depends on modeling choices that are hard to discipline empirically, like the assumed set of risk-sharing mechanisms available to households, the nature of their idiosyncratic risk, and the timing and distribution of fiscal transfers.[1]

In this paper, we propose a method to quantify the importance of imperfect risk sharing for the business cycle and to help researchers discipline these modeling choices. We extend a result by Nakajima (2005) and show that the aggregate implications of a large class of economies with incomplete markets can be equivalently studied in a representative-agent economy with wedges. We measure these wedges using households-level data and combine them with the representative-agent model to evaluate the contribution of imperfect risk sharing for the U.S. business cycle. We find that these frictions account for only 7% of output volatility on average, but they can have much larger effects when nominal interest rates are at the zero lower bound. Indeed, we find that these frictions were key determinants of the depth and persistence of the Great Recession.

We apply our methodology to a class of New Keynesian models with heterogeneous agents. To keep the analysis transparent, the "macro block"—the details of production, nominal rigidities, the conduct of monetary policy, etc.—of these economies is the same as in the standard three equations model. The households' decision problem, or "micro block", is instead modeled in a more flexible way: households face idiosyncratic risk, and their ability to smooth these shocks depends on the risk-sharing mechanisms available to them, e.g. the set of financial assets they can trade, their financial constraints and the

[1] A prominent example is Kaplan and Violante (2014), which shows that the consumption response to fiscal transfers is very different if households can trade one liquid asset or one liquid and one illiquid asset. Other modeling choices, which are inconsequential in representative agent economy, matter in heterogeneous agent economies such as the timing and distribution of the fiscal transfers (Kaplan, Moll, and Violante, 2018), how profits get distributed across households (Broer, Hansen, Krusell, and Oberg, 2018), and the cyclicality of idiosyncratic risk and access to liquidity (Werning, 2015).

presence of redistributive fiscal transfers. We formulate these features so as to nest most of the specifications considered in the literature including incomplete market models in the Bewley-Hugget-Aiyagari tradition as well as those with endogenous debt limits as in Kehoe and Levine (1993) and Alvarez and Jermann (2000).

Our analysis builds on an equivalence result between these economies with heterogeneous agents and the canonical representative-agent New Keynesian model. For any heterogeneous agent economy in our class, the macroeconomic aggregates—output, inflation and nominal interest rates—satisfy the equilibrium conditions of a representative-agent economy with two wedges: one affecting the discount factor of the stand-in household and one affecting her labor supply. We call this the *RA representation*.

The discount factor wedge captures the failure of aggregation when consumption risk is not perfectly shared across households. To explain why the RA representation features this wedge, consider an household that is on her Euler equation in an heterogeneous agent economy. With isoelastic preferences and complete markets, her consumption is a constant fraction of aggregate consumption at every point in time. Thus, the Euler equation also holds for aggregate consumption, and there is no need for a discount factor wedge in the RA representation. If risk sharing is not perfect, instead, consumption inequality can vary over time and households' consumption is no longer proportional to that of the aggregate. Therefore, the Euler equation evaluated using aggregate consumption does not typically hold, and the RA representation features a discount factor wedge.

The RA representation also features a wedge in the labor supply condition because of compositional changes in hours worked. To understand this result, consider an economy in our class and suppose that there is an increase in the cross-sectional dispersion of labor productivity. Because of substitution effects, high-productivity households will work more and low-productivity households will work less, a change in the composition of the labor force that leads to an increase in aggregate hours when measured in efficiency units. These compositional changes in worked hours that occur in the heterogeneous agent economy are captured in the RA representation by a wedge in the labor supply condition.

We derive expressions for these wedges as functions of households' consumption shares and relative wages and emphasize two key properties. First, the mapping between the wedges and the micro allocation is the same for every model in our class. This implies that we do not need to take a stand on the details of the micro block for the purpose of measuring the wedges, as long as we have observations on households' consumption and wages. Second, the wedges are a sufficient statistic for how households' heterogeneity affects aggregate variables, in the sense that shocks and frictions at the micro level matter for the aggregate if and only if they generate time variation in these two statistics. This

makes the wedges ideal empirical targets for the analysis of incomplete market economies.

We use panel data from the Consumer Expenditure Survey to measure the two wedges for the U.S. economy over the 1992-2017 period. The labor supply wedge does not display much variation at business cycle frequencies, and it mostly reflects the secular increase in labor income inequality over this period. The discount factor wedge is, instead, countercyclical and it displays a persistent increase after the Great Recession.

We then turn to study the aggregate implications of these movements using the RA representation. It is well known that an increase in the discount factor can induce sizable output drops in representative-agent New Keynesian models, especially when the zero lower bound constraint binds (Christiano, Eichenbaum, and Rebelo, 2011).[2] Indeed, these models typically need large increases in the discount factor to explain episodes of low interest rates, inflation and output, as for example the U.S. Great Recession. Frictions impeding risk sharing may be the root cause of these dynamics, and the increase in the discount factor wedge that we document provides *qualitatively* some support to this view. An important question is whether these movements are large enough to be *quantitatively* important.

To address this question, we use the realization of the wedges and data on output, inflation and nominal interest rates to jointly estimate the structural parameters of the RA representation and the stochastic process of the wedges. Given the estimated parameters, we use the RA representation to construct the counterfactual path for aggregate variables in an economy with complete financial markets—that is, an economy with time-invariant consumption shares. We show that the presence of complete financial markets reduces the standard deviation of output by only 7%, suggesting that deviations from perfect risk sharing contribute little on average to business cycle fluctuations. This is because, under the estimated monetary policy rule, increases in the discount factor wedge are offset by a reduction in nominal interest rates, with little effects on aggregate demand.

To further explore the role of the policy rule, we perform the same counterfactual during the Great Recession, an episode where the monetary authority was constrained by the zero lower bound. In this case, we find that imperfect risk sharing explains a third of the observed output losses and it helps accounting for the slow recovery. This result underscores the importance of accounting for constraints on monetary policy when evaluating

[2]This is especially true for New Keynesian models that, unlike the one we study here, feature capital accumulation. Away from the zero lower bound, an increase in the discount factor typically generates a comovement problem between consumption and investment, as first suggested by Barro and King (1984) for neoclassical models. At the zero lower bound, this does not happen because the higher discount factor can lead to higher real interest rates.

the aggregate implications of imperfect risk sharing.[3]

We conclude our paper by studying what feature of the micro data is responsible for the rise in the discount factor wedge during the Great Recession and what this tells us about the economic mechanisms behind our results. In our application, the discount factor wedge is the sample average of the inverse change in the consumption shares for a group of households that we identify as financially unconstrained. We show that the increase in this statistic is driven mostly by an increase in the dispersion of consumption growth for these households. This pattern cannot be rationalized by models with simple form of heterogeneity, such as the "two-agent" New Keynesian model studied in Galí, López-Salido, and Vallés (2007) and Bilbiie (2008). However, it is consistent with models in the Bewley-Hugget-Ayiagari tradition, specifically those that have emphasized heightened consumption risk and precautionary savings as a key driver of the fall in aggregate consumption during the Great Recession.

The economic literature has proposed two main mechanisms that can explain the increase in consumption risk of unconstrained households: higher volatility of their labor income, as for example in Bayer, Lütticke, Pham-Dao, and Tjaden (2019), versus a deterioration of the risk sharing mechanisms available to them, such as a tightening of credit constraints as in Guerrieri and Lorenzoni (2017). To explore which of these two mechanisms better accounts for the observed patterns, we study the relation between consumption and income during the Great Recession. We find that the data favors the second explanation, as the documented increase in the dispersion of consumption growth for unconstrained households is mostly due to higher sensitivity of consumption to income rather than an increase in the dispersion of the latter.

Related Literature. Our research contributes to a growing literature that introduces heterogeneous agents and incomplete financial markets in New Keynesian models. Thanks to recent computational advances, researchers have used these environments to study how frictions impeding risk sharing across households affect the transmission mechanism of monetary and fiscal policy and more generally the business cycle.[4] Closely related to our work are the papers of Bayer, Born, and Luetticke (2020), Auclert, Rognlie, and Straub

[3]See also Schaab (2020) for a discussion of this point in an heterogeneous agent New Keynesian model with a zero lower bound constraint.

[4]See Kaplan, Moll, and Violante (2018); Auclert (2017); McKay, Nakamura, and Steinsson (2016); McKay and Reis (2016); Hagedorn, Manovskii, and Mitman (2019); Gornemann, Kuester, and Nakajima (2016); Bhandari, Evans, Golosov, and Sargent (2018) for recent contributions on monetary and fiscal policy. Guerrieri and Lorenzoni (2017) and Jones, Midrigan, and Philippon (2018) study how credit constraints can affect aggregate demand, while McKay (2017), Challe, Matheron, Ragot, and Rubio-Ramirez (2017), Den Haan, Rendahl, and Riegler (2017), Heathcote and Perri (2018), Ravn and Sterk (2017) and Bayer et al. (2019) focus on idiosyncratic income risk. See Krueger, Mitman, and Perri (2016) and Kaplan and Violante (2018) for literature reviews.

(2020) and Bilbiie, Primiceri, and Tambalotti (2022), who estimate medium sized New Keynesian models with heterogeneous agents. These are fully structural models and can be used to perform a variety of counterfactuals. However, they require the researchers to specify the details of the micro block—such as the set of financial assets available to households, their borrowing constraints, the risk they face, etc.

The main contribution of our paper is to recognize that we can be agnostic about these details of the micro block when performing *some* of these counterfactuals, as long as we observe households' consumption choices. Specifically, our methodology allows one to assess the business cycle effects of imperfect risk sharing. We think that this is important for two reasons. First, our approach is less subject to misspecifications of the micro block and, by nesting most of the frameworks studied in this literature, it provides a benchmark calculation. Second, our paper identifies two sufficient statistics for the aggregate implications of households' heterogeneity in a broad class of models, and it suggests an approach to measure them. In this sense, our contribution is complementary to structural modeling because researchers could fruitfully use these wedges as a calibration target to discipline their models. We believe this is important because, unlike structural modeling, our approach cannot be used for welfare assessment or policy analysis.[5]

The counterfactuals that we perform are related to the business cycle accounting methodology of Chari, Kehoe, and McGrattan (2007). There are two main differences between these procedures. First, in our approach the wedges are measured using household-level observations, rather than being chosen to replicate the observed path of aggregate data. By doing so, we are able to isolate the wedges due to imperfect risk sharing.[6] Second, our main quantitative experiment constructs the path for macroeconomic variables in a counterfactual economy with complete financial markets. This is not equivalent to the approach of Chari, Kehoe, and McGrattan (2007), which assesses the effects of specific wedges on the business cycle.

The idea that heterogeneous agents' economies admit an RA representation with wedges was developed by Maliar and Maliar (2001, 2003) for complete markets economies following the insight of Constantinides (1982), and by Nakajima (2005) and more recently by Werning (2015) and Debortoli and Galí (2017, 2022) for economies with incomplete markets. In general, these time-varying wedges are endogenous equilibrium objects in the hetero-

[5]The latter is due to the endogeneity of the wedges to the policy regime. In this sense, our methodology is valid only for *ex-post evaluations* similar to Arkolakis, Costinot, and Rodríguez-Clare (2012) in the international trade literature.

[6]For example, if we were to only use aggregate data to measure the wedge in the Euler equation of the RA representation we would not be able to distinguish the discount factor wedge due to imperfect risk sharing from any other friction that would take the form of an Euler equation wedge. By using micro-data, we are able to separately identify the former.

geneous agent economy under consideration, although recent papers derive the mapping between primitives and the wedges in some specific economies (Werning, 2015; Acharya and Dogra, 2018). Krueger and Lustig (2010) provide conditions under which the discount factor wedge is constant over time and it is therefore irrelevant for aggregate fluctuations. Our paper is the first to exploit this representation to quantify the macroeconomic implications of imperfect risk sharing.

Finally, there is a connection between our paper and the literature that evaluates the asset pricing implications of models with heterogeneous households. See for example Brav, Constantinides, and Geczy (2002), Vissing-Jørgensen (2002), Krueger, Lustig, and Perri (2008), and Kocherlakota and Pistaferri (2009). The goal of these papers is to estimate the stochastic discount factor with micro data given a particular form of market incompleteness. This is similar to the construction of the discount factor wedge in our approach. Clearly the scope of our analysis differs from these papers.

Layout. The paper is organized as follows. Section 2 introduces the class of heterogeneous agents economies at the center of our application. Section 3 derives the RA representation and discusses why this representation is a useful tool for the evaluation of heterogeneous agents models. Section 4 discusses how we estimate the preference wedges using panel data and approximate their stochastic process. In Section 5 we measure the preference wedges and combine these series with the RA representation to measure the role of imperfect risk sharing for the U.S. business cycle. We finish this section by discussing which models are most consistent with the patterns we identify in the micro data. Section 6 concludes.

2 New Keynesian models with heterogeneous agents

In this section we introduce a class of New Keynesian models with heterogeneous households. The models in this class share the same specification for households' preferences, technology, market structure and the conduct of monetary policy—elements that are borrowed from the prototypical "three equations" New Keynesian framework. However, they can differ in the details of the households' decision problem, such as the cyclicality of idiosyncratic risk faced by households, the set of assets they can trade, their financial constraints, as well as the timing and distribution of fiscal transfers.

Environment. Time is discrete and indexed by $t = 0, 1, ...$ The economy is populated by a continuum of households, final good producers, intermediate good firms, and the

monetary authority. There are two types of states: aggregate and idiosyncratic. We denote the aggregate state by z_t and the idiosyncratic state by v_t, both of which are potentially vector valued. Let $z^t = (z_0, z_1, ..., z_t)$ be a history of realized aggregate states up to period t and $v^t = (v_0, v_1, ..., v_t)$ be a history of idiosyncratic states up to period t. We also let $s_t = (z_t, v_t)$ and $s^t = \left(z^t, v^t\right)$. Let $\Pr\left(s^t|s^{t-1}\right)$ be the probability of a history s^t given s^{t-1}. We assume that $\Pr(s^t|s^{t-1}) = \Pr(v^t|v^{t-1}, z^t)\Pr(z^t|z^{t-1})$ and allow for the possibility that the aggregate states affect the distribution of the idiosyncratic states. To reduce the notation, the initial idiosyncratic state v_0 also indexes the initial distribution of assets for such agent and the initial aggregate state z_0 indexes the initial distribution of such variables. Thus, without loss of generality, we express all the individual variables as a function of history s^t and, in a symmetric equilibrium, aggregate quantities and prices are functions of z^t.

Households are infinitely lived and have preferences over consumption, $c\left(s^t\right)$, and hours worked, $l\left(s^t\right)$, given by

$$\sum_{t=0}^{\infty}\sum_{s^t}\beta^t \Pr\left(s^t|s_0\right)\tilde{\theta}(z^t)U\left(c\left(s^t\right), l\left(s^t\right)\right), \tag{1}$$

where β is the discount factor and $\tilde{\theta}(z^t)$ is a shock to the marginal utility of consumption and disutility of labor defined recursively as $\tilde{\theta}(z^{t+1}) = \theta(z_{t+1})\tilde{\theta}(z^t)$. We further assume that the period utility is given by

$$U(c, l) = \frac{c^{1-\sigma} - 1}{1-\sigma} - \chi\frac{l^{1+\psi}}{1+\psi}, \tag{2}$$

with $\sigma > 0$ and $\psi > 0$.

The final good is produced combining differentiated intermediate goods according to the technology

$$Y(z^t) = \left(\int_0^1 y_j(z^t)^{\frac{1}{\mu}}dj\right)^{\mu}, \tag{3}$$

where μ is related to the (constant) elasticity of substitution across varieties, ε, by the following, $\mu = \varepsilon/(\varepsilon - 1)$. The intermediate inputs indexed by j are produced using labor

$$y_j(z^t) = A(z_t)n_j(z^t), \tag{4}$$

where $A(z_t)$ is an aggregate technology shock, common across firms, and $n_j(z^t)$ is labor in *efficiency units* utilized by the producer of intermediate good j. Feasibility requires that

$$\int n_j(z^t)dj = \sum_{v^t}\Pr(v^t|z^t)e(v_t)l(v^t, z^t), \tag{5}$$

where the right side is the supply of labor in efficiency units. Each individual v^t is associated to a particular level of efficiency $e(v_t)$: hiring more high-efficiency types, holding total hours worked fixed, results in higher output produced by the firm. This individual-specific productivity shock $e(v_t)$ generates idiosyncratic labor income risk for households.

We now describe the market structure for this economy with a particular emphasis on the households side.

Households. Households enter the period with financial assets and they work for intermediate good producers. They choose consumption, new financial positions and labor in order to maximize their expected life-time utility.

We model financial markets in a flexible way. Households can trade a risk-free nominal bond. We denote by $b(s^t)$ the position taken today by a household and by $1 + i\left(z^t\right)$ the nominal return on the bond. Households can also trade a set $\mathcal{K}$ of additional assets, with the nominal payout of a generic asset $k \in \mathcal{K}$ given by $R_k(s^t, s_{t+1})$. We let $q_k(s^t)$ be the price of the asset. This formulation allows for different types of financial assets: individual Arrow securities, shares of the intermediate good firms, complex financial derivatives, etc. We let $a_k\left(s^{t-1}\right)$ be the holdings of assets k that a household with history v^{t-1} has accumulated after an aggregate history z^{t-1}. Trades in these additional financial assets potentially require transaction costs $\mathcal{T}(\{a_k(s^{t-1})\}_{k\in\mathcal{K}}, \{a_k(s^t)\}_{k\in\mathcal{K}}, s^t)$ that can depend on the inherited portfolio $\{a_k(s^{t-1})\}_{k\in\mathcal{K}}$, the new portfolio $\{a_k(s^t)\}_{k\in\mathcal{K}}$, and s^t. The transaction costs do not apply to $b(s^t)$, so our framework does not nest limited participation economies where agents must pay a fixed cost to change their position in nominal bonds.

In addition, we allow for a number of constraints that potentially restricts the financial positions that households can choose,

$$H\left(b\left(s^t\right), \left\{a_k\left(s^t\right)\right\}_{k\in\mathcal{K}}, s^t\right) \geq 0 \tag{6}$$

for some vector-valued function H. We refer to the set of constraints in (6) as *trading restrictions*. We assume that purchasing risk-free nominal bonds weakly relaxes the trading restrictions, $\partial H\left(b, \{a_k\}_{k\in\mathcal{K}}, s^t\right)/\partial b \geq 0$.

The set of assets $\mathcal{K}$, the transaction costs $\mathcal{T}$, and the trading restrictions in (6) are a flexible way of representing different sets of risk-sharing mechanisms available to households. Our formulation nests the economy with complete financial markets, when the set of assets spans all possible aggregate and idiosyncratic histories and there are no transaction costs or trading restrictions. In addition, it encompasses as special cases a large class of models with incomplete financial markets: the Bewley-Huggett-Aiyagari economy, the two-assets

economy in Kaplan and Violante (2014) and Kaplan, Moll, and Violante (2018), the endogenous debt limits in Alvarez and Jermann (2000), or the various restrictions on asset trading in Chien, Cole, and Lustig (2011, 2012). Note, also, that the H function can depend on s^t, so we allow for aggregate and idiosyncratic shocks to affect households' financial constraints. Moreover, our formulation allows for heterogeneity in households' access to assets other than the risk-free nominal bond, a property that is critical to account for the wealth distribution in the data.

Given initial assets' holdings, households choose $\{c(s^t), l(s^t), b(s^t), \{a_k(s^t)\}_{k\in\mathcal{K}}\}$ to maximize their utility subject to the trading restrictions in (6) and the nominal budget constraint,

$$\begin{aligned} P\left(z^t\right) c\left(s^t\right) + \frac{b\left(s^t\right)}{1+i\left(z^t\right)} + \sum_{k\in\mathcal{K}} q_k(s^t) a_k\left(s^t\right) + P(z^t)\mathcal{T}(\{a_k(s^{t-1})\}_{k\in\mathcal{K}}, \{a_k(s^t)\}_{k\in\mathcal{K}}, s^t) \\ \leq W\left(z^t\right) e(v_t) l\left(v^t, z^t\right) - T(s^t) + b\left(s^{t-1}\right) + \sum_{k\in\mathcal{K}} R_k\left(s^{t-1}, s_t\right) a_k\left(s^{t-1}\right), \end{aligned}$$

where $W\left(z^t\right)$ is the nominal wage per efficiency units and $T(s^t)$ are lump-sum taxes.

Because of the assumption that $\partial H/\partial b \geq 0$, a necessary condition for optimality is

$$\frac{1}{1+i\left(z^t\right)} \geq \beta \sum_{s_{t+1}} \left\{ \frac{\Pr\left(s^{t+1}|s^t\right)\theta(z_{t+1})}{1+\pi\left(z^{t+1}\right)} \left[\frac{c\left(s^t, s_{t+1}\right)}{c\left(s^t\right)}\right]^{-\sigma} \right\}, \tag{7}$$

where $\pi(z^{t+1}) = P(z^{t+1})/P(z^t) - 1$ is the net inflation rate. This condition must hold with equality if the trading restrictions on the nominal bond do not bind. For the rest of the paper, we assume that there always exist an agent for which equation (7) holds as an equality. Because $\partial H/\partial b \geq 0$, equation (7) holds with equality for households with the highest valuation for the risk-free bond.

The condition for the optimality of labor supply is

$$\chi l(s^t)^{\psi} = w(z^t) e(v_t) c(s^t)^{-\sigma} \tag{8}$$

where $w(z^t) = W(z^t)/P(z^t)$ is the real wage per efficiency unit.

Final good producers. The final good is produced by competitive firms that operate the production function in (3). From their decision problem, we can derive the demand function for the j-th variety

$$y_j(z^t) = \left(\frac{P_j\left(z^t\right)}{P(z^t)}\right)^{\mu/(1-\mu)} Y(z^t) \tag{9}$$

where $P_j(z^t)$ is the price of variety j and $P(z^t) = \left[\int P_j(z^t)^{1/(1-\mu)} dj\right]^{1-\mu}$ is the price index.

Intermediate good producers. Each intermediate good is supplied by a monopolistic competitive firm. The monopolist of variety j operates the technology (4). The firm faces quadratic costs to adjust their prices relative to the inflation target of the monetary authority $\bar{\pi}$,

$$\frac{\kappa}{2}\left[\frac{P_j\left(z^t\right)}{P_j\left(z^{t-1}\right)\left(1+\bar{\pi}\right)} - 1\right]^2. \tag{10}$$

The problem of firm j is to choose its price $P_j\left(z^t\right)$ given its previous price $P_j\left(z^{t-1}\right)$ to maximize the present discounted value of real profits. As is well known, state prices in incomplete market economies are not uniquely determined. The issue is even more complex in our framework because we are purposefully not fully specifying the set of assets available and the trading restrictions in (6). We resolve this indeterminacy by assuming that the firm discounts future profits using the real state price

$$Q\left(z^{t+1}\right) = \beta \max_{v^t}\left\{\Pr\left(z^{t+1}|z^t\right)\theta(z_{t+1})\sum_{v_{t+1}}\Pr\left(v^{t+1}|z^{t+1},v^t\right)\left[\frac{c\left(z^{t+1},v^{t+1}\right)}{c\left(z^t,v^t\right)}\right]^{-\sigma}\right\}. \tag{11}$$

That is, firms discount future profits using the marginal rate of substitution of the agent that values dividends in a given aggregate state next period the most. This would be the equilibrium state price if all agents could trade aggregate Arrow securities.

The firm's problem can be written recursively as

$$V\left(P_j,z^t\right) = \max_{p_j,y_j,n_j}\frac{p_j y_j}{P\left(z^t\right)} - w(z^t)n_j(z^t) - \frac{\kappa}{2}\left[\frac{p_j}{P_j(1+\bar{\pi})} - 1\right]^2 + \sum_{z^{t+1}} Q(z^{t+1}|z^t)V\left(p_j,z^{t+1}\right) \tag{12}$$

subject to the production function (4) and the demand function (9).

The solution to the firm's problem together with symmetry across firms requires that the following version of the New Keynesian Phillips curve holds in equilibrium

$$\tilde{\pi}\left(z^t\right) = \frac{1}{\kappa\left(\mu-1\right)}Y(z^t)\left[\mu\frac{w(z^t)}{A\left(z_t\right)} - 1\right] + \sum_{z^{t+1}} Q(z^{t+1}|z^t)\tilde{\pi}\left(z^{t+1}\right) \tag{13}$$

where we define $\tilde{\pi}\left(z^t\right) = [(\pi(z^t)-\bar{\pi})/(1+\bar{\pi})] \times [(\pi(z^t)+1)/(1+\bar{\pi})]$ and $w(z^t)/A(z_t)$ is the real marginal cost for producing a unit of the final good.

Monetary and fiscal policy. The monetary authority follows a standard Taylor rule

$$1+i(z^t)=\max\left\{[1+i(z^{t-1})]^{\rho_i}\left[(1+\bar{i})\left(\frac{1+\pi(z^t)}{1+\bar{\pi}}\right)^{\gamma_\pi}\right]^{1-\rho_i}\exp\{\varepsilon_m(z_t)\},1\right\}, \tag{14}$$

where $(1+\bar{i})=(1+\bar{\pi})/\beta$ is the nominal interest in a deterministic steady state of the model and $\varepsilon_m(z_t)$ is a monetary shock. Because of the zero lower bound, whenever the interest rate predicted by the Taylor rule is negative, the monetary authority sets the nominal interest rate to zero.

The evolution of the aggregate supply of the nominal bond, $B(z^t)$, and taxes, $T(s^t)$, must satisfy the government budget constraint,

$$B(z^{t-1})=\frac{B(z^t)}{1+i(z^t)}+\sum_{v^t}\Pr(v^t|z^t)T(z^t,v^t). \tag{15}$$

Equilibrium. In equilibrium, the labor market, goods markets, and financial markets clear. Specifically, market clearing in the nominal bond market requires that

$$\sum_{v^t}\Pr\left(v^t|z^t\right)b\left(z^t,v^t\right)=B(z^t). \tag{16}$$

For the other assets, given the asset supply $\bar{a}_k(z^t)$, market clearing requires that

$$\sum_{v^t}\Pr\left(v^t|z^t\right)a_k\left(z^t,v^t\right)=\bar{a}_k(z^t). \tag{17}$$

Moreover, since firms' equity is the only asset in positive net supply other than the nominal risk-free bond, the supply of the other assets available and their returns R_k must satisfy the following restriction to ensure that agents budget constraints is consistent with the aggregate resource constraint:

$$\sum_{k\in\mathcal{K}}\sum_{v^t}\Pr(v^t|z^t)R_k(z^t,v^t)\bar{a}_k(z^{t-1})-\sum_{k\in\mathcal{K}}q_k(z^t)\bar{a}_k(z^t)=D(z^t) \tag{18}$$

where $D(z^t)=[P(z^t)-W(z^t)/A(z_t)]Y(z^t)-\kappa[(1+\pi(z^t))/(1+\bar{\pi})-1]^2/2$ are aggregate nominal firm profits.

We can then define an equilibrium for this economy.

Definition 1. *Given an asset structure $(\mathcal{K},\mathcal{T},R_k,\bar{a}_k,H)$, the distribution of initial assets and lagged prices, an equilibrium is a set of households' allocations $\{c(s^t),l(s^t),b(s^t),a_k(s^t)\}$, a fiscal*

policy $\{B(z^t), T(s^t)\}$, *prices* $\{P(z^t), W(z^t), 1+i(z^t), Q(z^t), q_k(z^t)\}$, *and aggregates* $\{C(z^t), Y(z^t)\}$ *such that i) the households' allocation solves the households' decision problem, ii) the price for the final good solves (12) with* $P(z^t) = P_j(z^t)$, *iii) the state price is given by (11), iv) the nominal interest rate satisfies the Taylor rule (14), v) the government budget constraint (15) is satisfied, and vi) markets clear in that (16)–(18) hold and*

$$Y(z^t) = C(z^t) + \frac{\kappa}{2}\left[\frac{\pi(z^t) - \bar{\pi}}{1+\bar{\pi}}\right]^2 + \mathcal{T}(z^t)$$

where aggregates are given by

$$\begin{aligned} Y(z^t) &= A(z_t)\sum_{v^t} Pr(v^t|z^t)e(v_t)l(v^t, z^t), \\ C(z^t) &= \sum_{v^t} \Pr(v^t|z^t)c(z^t, v^t), \end{aligned}$$

and $\mathcal{T}(z^t)$ *are the aggregate transaction costs,*

$$\mathcal{T}(z^t) = \sum_{v^t} \Pr(v^t|z^t)\mathcal{T}(\{a_k(s^{t-1})\}_{k\in\mathcal{K}}, \{a_k(s^t)\}_{k\in\mathcal{K}}, s^t).$$

3 The RA representation

We now show that the aggregate variables in the class of New Keynesian models just described can be equivalently derived from the equilibrium conditions of a representative agent economy with wedges. We refer to this as the *RA representation*. Section 3.1 derives the RA representation, while Section 3.2 discusses how we can use it for counterfactual analysis. Section 3.3 and 3.4 study specific economies nested in our framework that admits an analytical mapping between the model primitives and the wedges. This analysis will be useful to build intuition on how different types of shocks and frictions studied in the literature affect the wedges on the RA representation.

3.1 Equilibrium representation

Toward establishing the RA representation for our class of New Keynesian models, let us define the following statistics:

$$\beta(v^t, z^{t+1}) \equiv \sum_{v_{t+1}} \Pr\left(v_{t+1}|v^t, z^{t+1}\right) \left(\frac{c(z^{t+1}, v^t, v_{t+1})/C(z^{t+1})}{c(z^t, v^t)/C(z^t)}\right)^{-\sigma} \tag{19}$$

$$\omega\left(z^t\right) \equiv \left[\sum_{v^t} \Pr\left(v^t|z^t\right) \left(\frac{c(z^t, v^t)}{C(z^t)}\right)^{\frac{-\sigma}{\psi}} e\left(v_t\right)^{\frac{1+\psi}{\psi}}\right]^{-\psi}. \tag{20}$$

We then have the following proposition where we assume that the aggregate transaction costs, $\mathcal{T}(z^t)$, are negligible.

Proposition 1. *Suppose that* $\left\{C\left(z^t\right), Y\left(z^t\right), \pi\left(z^t\right), i(z^t), Q(z^{t+1})\right\}$ *are part of an equilibrium of an heterogeneous agent economy described in Section 2. Then, they must satisfy the aggregate Euler equation,*

$$\frac{1}{1+i\left(z^t\right)} = \beta \max_{v^t} \sum_{z_{t+1}} \Pr\left(z^{t+1}|z^t\right) \left[\frac{\theta(z_{t+1})\beta\left(v^t, z^{t+1}\right)}{1+\pi\left(z^{t+1}\right)} \left(\frac{C\left(z^{t+1}\right)}{C\left(z^t\right)}\right)^{-\sigma}\right], \tag{21}$$

the Phillips curve,

$$\tilde{\pi}\left(z^t\right) = \frac{Y(z^t)}{\kappa\left(\mu-1\right)} \left[\mu \frac{\chi Y\left(z^t\right)^{\psi} C\left(z^t\right)^{\sigma} \omega\left(z^t\right)}{A(z_t)^{1+\psi}} - 1\right] + \sum_{z^{t+1}} Q(z^{t+1}|z^t)\tilde{\pi}\left(z^{t+1}\right) \tag{22}$$

the Taylor rule (14), the resource constraint,

$$Y(z^t) = C(z^t) + \frac{\kappa}{2}\left[\frac{\pi(z^t) - \bar{\pi}}{1+\bar{\pi}}\right]^2, \tag{23}$$

and

$$Q\left(z^{t+1}|z^t\right) = \beta \max_{v^t} \left\{\beta\left(v^t, z^{t+1}\right) \Pr\left(z^{t+1}|z^t\right) \theta(z_{t+1}) \left(\frac{C\left(z^{t+1}\right)}{C\left(z^t\right)}\right)^{\sigma}\right\}, \tag{24}$$

given $\{\beta(v^t, z^{t+1}), \omega(z^t)\}$ *defined in (19) and (20).*

The proof for this result is straightforward, and it extends the result of Nakajima (2005) to an economy with nominal rigidities.

The aggregate Euler equation (21) is obtained by substituting the households' marginal rate of substitution in (7) using the definition of $\beta(v^t, z^{t+1})$ and noting that under our

assumptions the Euler equation holds for agents with the highest valuation of the bond—those attaining the maximum in equation (21).

For the aggregate Phillips curve (22) we proceed as follows. We raise both sides of households' labor supply condition (8) by $1/\psi$, multiply them by $e(v_t)C(z^t)^{\sigma/\psi}$, and average across households to obtain

$$\chi^{\frac{1}{\psi}}\left[\sum_{v^t}\Pr(v^t|z^t)e(v_t)l(s^t)\right]C(z^t)^{\frac{\sigma}{\psi}}=w(z^t)^{\frac{1}{\psi}}\left[\sum_{v^t}\Pr\left(v^t|z^t\right)\left(\frac{c(z^t,v^t)}{C(z^t)}\right)^{\frac{-\sigma}{\psi}}e\left(v_t\right)^{\frac{1+\psi}{\psi}}\right].$$

We can then use the production function (4) and the definition of $\omega(z^t)$ in (20) to express the real wage as

$$w(z^t)=\chi\left[\frac{Y(z^t)}{A(z_t)}\right]^{\psi}C(z^t)^{\sigma}\omega(z^t).$$

We then substitute the above expression in equation (13) to obtain the aggregate Phillips curve (22).

The equilibrium conditions (14), (21), (22) and (23) are equivalent to those of a representative agent New Keynesian with two wedges: one in the Euler equation and one in the Phillips curve. We refer to the first as the discount factor wedges, and to the second as the labor supply wedge. These wedges are functions of the individual allocations in the original heterogeneous agent economy and do not typically have a structural interpretation.

Discussion. Before moving forward, we discuss some potential limitations of our framework. First, while we have been flexible on the households' decision problem, we made restrictive assumptions about other aspects of the model. For instance, there is no capital accumulation in this economy, wages are perfectly flexible, all movements in labor are at the intensive margin, and we have taken a stand on some of the aggregate shocks driving the economy—the technology, monetary and preference shock. In Online Appendix A we show that it is relatively straightforward to derive an equivalent of Proposition 1 for economies with a more complex macro block. Specifically, we derive the RA representation in a model that features capital accumulation, one with frictions in financial intermediation as in Gertler and Karadi (2011) and Gertler and Kiyotaki (2010), and for a small open economy model where asset prices matter for the amount of borrowing as in Mendoza (2010). All these economies admit an RA representation similar to the one in Proposition 1, with a discount factor and labor supply wedge as defined in (19) and (20). Intuitively, this is because market incompleteness in these economies matters for firms' production choices

only through its impact on their stochastic discount factor.[7] Thus, these model ingredients do not affect the mapping between the micro-level allocations and the wedges, but they entail a different propagation of these wedges to aggregate outcomes.

Second, we have restricted households' preferences to be separable and isoelastic. While the definition of the wedges relies on this assumption, it is worth pointing out that we can obtain similar RA representations for different sets of preferences. For example, in Online Appendix C.1 we consider an economy where households' preferences allow for corner solutions in worked hours. There, we show that this economy admits an RA representation as in our benchmark economy, but with a slightly different labor supply wedge. In addition, Online Appendix A.4 studies an economy where households have preferences over durable and non-durable consumption goods. Also in this case we derive the RA representation and find it to be quite similar to the one of our benchmark. Interestingly, the two are identical when the underlying heterogeneous agents economy features no dispersion across households in the ratio between durable and non-durable consumption.

Third, we have assumed that firms discount future profits using the marginal rate of substitution of the agent that values dividends the most.[8] This choice is arbitrary, but it does not affect the definition of the wedges in the RA representation. That is, different assumptions on firms capital structure choices and dividend distribution policies affect households' consumption, but do not alter the mapping between households consumption choices and the wedges.

3.2 Using the RA representation for model evaluation

Proposition 1 has two main implications. The first implication is that the wedges summarize all the information from the micro block of the model that is needed to characterize the behavior of aggregate variables. That is, they are *sufficient statistics* for the specifics of the model regarding the set of assets traded, the transaction costs and trading restrictions faced by households, the fiscal policy $\{B(z^t), T(s^t)\}$, and the nature of their idiosyncratic risk. The second implication is that the mapping between individual allocations and the wedges is *invariant* to the details of the micro block, in the sense that for all the economies nested in the environment of Section 2, the relation between $\{\beta(v^t, z^{t+1}), \omega(z^t)\}$ and the

[7] This property does not hold in economies where the joint distribution of capital and entrepreneurial ability directly matters for production choices, as for instance in Buera and Moll (2015). Differently from our framework, the RA representation of these economies will also feature an "efficiency wedge" in the terminology of Chari, Kehoe, and McGrattan (2007).

[8] An alternative would be to follow the approach in Makowski (1983a) and Makowski (1983b) and let the state price used by the firm be the valuation of the agent that maximizes the firm's stock value for any possible choices for the firm, assuming that agents can trade stocks without frictions. See Bisin, Clementi, and Gottardi (2017) for a recent illustration of the appealing implications of this approach.

households' allocation is the same and it is given by equation (19) and (20). These two properties, in turn, make the RA representation a useful device for the empirical analysis of New Keynesian economies with heterogeneous agents.

First, suppose that we have a procedure to measure $\{\beta(v^t, z^{t+1}), \omega(z^t)\}$ using observations on households' consumption shares and labor productivities. Because these are sufficient statistics for how the "micro block" affects macroeconomic variables, they also provide an informative empirical target for the calibration/evaluation of a specific model. Researchers that wish to use a specific model nested in this class should make sure that their model fits the statistical properties of the wedges observed in the data.

Second, we can use the RA representation to address a question that goes back to the seminal work of Krusell and Smith (1998)—to what extent imperfect risk sharing across households matters for aggregate fluctuations. To address this question, one needs to compare the behavior of observed macroeconomic variables to those that would arise in a world where households could perfectly insure their idiosyncratic risk. These counterfactuals can be computed using the RA representation, as the next result shows.

Proposition 2. *Consider the class of heterogeneous agent economies described in Section 2 and suppose that financial markets are complete—the set of assets $\mathcal{K}$ contains Arrow securities contingent on the realizations of the aggregate and idiosyncratic state and there are no trading restrictions other than a non-binding no-Ponzi condition. Then, if $\{C^{cm}(z^t), Y^{cm}(z^t), \pi^{cm}(z^t), i^{cm}(z^t), Q^{cm}(z^{t+1})\}$ are part of an equilibrium they satisfy (14), (21), (22), (23) and (24) with*

$$\beta^{cm}\left(v^t, z^{t+1}\right) = 1 \tag{25}$$

$$\omega^{cm}(z^t) = \left[\sum_{v^t} Pr(v^t|z^t)\varphi(v_0)^{\frac{-\sigma}{\psi}} e(v_t)^{\frac{1+\psi}{\psi}}\right]^{-\psi}. \tag{26}$$

This result is an application of the aggregation result of Constantinides (1982) that Maliar and Maliar (2001, 2003) extend to economies with endogenous labor supply. Given isoelastic preferences, the complete market economy features constant consumption shares for households, implying that $\beta^{\text{cm}}\left(v^t, z^{t+1}\right) = 1$. So, the Euler equation of the heterogeneous agent economy with complete markets coincides with the Euler equation of the representative-agent economy. Aggregate labor supply in the complete-market economy, however, can differ from that of the representative-agent economy because of substitution effects due to idiosyncratic productivity shocks. These differences are captured by the wedge $\omega^{\text{cm}}(z^t)$.[9]

[9]To understand this expression, suppose that $\psi = 1$ and households have the same initial wealth, so that

To explain how we can leverage Proposition 1 and 2 to quantify the macroeconomic effects of imperfect risk sharing, let us assume for now that we know the probability distribution of z^t and the stochastic process for $\{\theta(z^t), A(z^t), \varepsilon_m(z^t), \beta(v^t, z^{t+1}), \omega(z^t), \omega^{\text{cm}}(z^t)\}$. Thus, given a realization of z^t, we can use the RA representation of Proposition 1 to obtain the underlying equilibrium path for aggregate variables—output, inflation and nominal interest rates. We label these the *actual paths*. We can then compare these paths to the *complete markets paths* computed by setting $\beta\left(v^t, z^{t+1}\right) = 1$ and $\omega(z^t) = \omega^{\text{cm}}(z^t)$ in the RA representation. These paths will differ because they feature a different realization and a different stochastic process for the wedges: by comparing the two we are able to isolate the impact that imperfect risk sharing has for macroeconomic aggregates over the particular history z^t.

In order to carry out this counterfactual, we need a procedure to measure the realization of $\{\beta(v^t, z^{t+1}), \omega(z^t), \omega^{\text{cm}}(z^t)\}$ using households' level data and to approximate their stochastic process. We will discuss these two issues in Section 4 and apply our framework to U.S. data in Section 5. Before moving there, though, let us discuss how different economic mechanisms maps in the wedges of the RA representation using some analytically tractable examples. We will focus mostly on the discount factor wedge because it turns out to be empirically the dominant factor in our application to the U.S. economy.

Readers less interested in these examples can move directly to Section 4 without losing the thread.

3.3 Households' precautionary savings and the discount factor wedge

We now illustrate Proposition 1 using two examples. In the first example we study an incomplete market economy where households' idiosyncratic income risk is time-varying. In this economy, an increase in the volatility of households' income leads to higher incentives to save for precautionary reasons, which can result in a fall in aggregate consumption and output—a mechanism studied recently by Bayer et al. (2019), Heathcote and Perri (2018) and Ravn and Sterk (2021) to explain the depth and persistence of the Great Recession. In the second example, instead, we consider an economy where households' precautionary saving motives are triggered by a time-varying borrowing constraint, as in Guerrieri and Lorenzoni (2017). In both examples, we show that these time-varying precautionary motives are captured in the RA representation by the discount factor wedge $\beta(v^t, z^{t+1})$.

$\varphi(v_0) = 1$. In this case, $\omega(z^t) = \text{Var}[e(v_t)|z^t]^{-1}$. When the variance of $e(v_t)$ increases, the labor supplied by high productivity households increases and the one supplied by the low productivity households declines because of a substitution effect. So, labor supply goes up when measured in efficiency units. This effect is captured by a decline in $\omega(z^t)$ in the RA representation—an increase in the labor supply of the stand-in household.

Labor income risk and the discount factor wedge. Let $\sigma = 1$ and the idiosyncratic productivity shocks evolve according to

$$\Delta \log[e(v_{t+1})] = -\frac{\sigma_e^2(z_t)}{2} + \varepsilon_{t+1} \qquad \varepsilon_{t+1}|z^{t+1} \sim \mathcal{N}\left(0, \sigma_e^2(z_t)\right).$$

That is, idiosyncratic productivity is a random walk with Gaussian shocks. The standard deviation of individual productivity growth varies over time with the aggregate state: when $\sigma_e^2(z_t)$ is high, households face higher income risk.

We assume that households can only trade the risk-free bonds in zero net supply and face the borrowing limit $b(s^t) \geq 0$.[10] These two assumptions imply that households cannot save in equilibrium.[11] Therefore, all households are hand-to-mouth and consume every period their after tax income, $e(v_t)\left[w(z^t)l(s^t) + T(z^t)\right]$. Furthermore, we can verify from the labor supply condition (8) and $\sigma = 1$ that $l(s^t)$ is the same across individuals. So, we have $c(s^t) = e(v_t)C(z^t)$ from the aggregate resource constraint.

Given the equilibrium consumption function, the relative marginal rate of substitution of the households are just functions of the idiosyncratic income process,

$$\left(\frac{c(z^{t+1}, v^t, v_{t+1})/C(z^{t+1})}{c(z^t, v^t)/C(z^t)}\right)^{-1} = \frac{e(v_t)}{e(v_{t+1})}.$$

Substituting these expressions in equation (19) and (20) we can compute the implied $\beta(v^t, z^{t+1})$ and $\omega(z^t)$ of this economy:

$$\begin{aligned} \beta(v^t, z^{t+1}) &= \exp\{\sigma_e^2(z_t)\} \\ \omega(z^t) &= 1. \end{aligned}$$

The discount factor wedge is the same across households and it depends on the realization of the aggregate shock at date t: the higher idiosyncratic income risk at date t, the more "patient" the stand-in household in the RA representation. The labor supply wedge is constant over time because households supply the same amount of labor in equilibrium.

This example is useful to understand how the interaction between idiosyncratic risk and incomplete financial markets can affect aggregate variables in New Keynesian models with incomplete markets. Suppose that households face higher idiosyncratic risk, that is

[10] Because households cannot trade stocks of the firms, we also assume that profits of the intermediate good producers are taxed by the government and rebated to households in proportion to the realization of idiosyncratic productivity.

[11] The literature refers to this example with tight borrowing limits and bonds in zero net supply as the *zero liquidity limit*. See Werning (2015) and Ravn and Sterk (2017) for examples.

$\sigma_e^2(z_t)$ increases. Because of incomplete markets, households respond to heightened risk by increasing their demand of the risk-free bond. In equilibrium, households are hand-to-mouth, so aggregate savings cannot change. Thus, the increase in households' demand of savings must be met in equilibrium by a fall in real interest rates and/or by a decline in their income at date t. Whether these precautionary motives are mostly reflected on prices or quantities depends on the response of the monetary authority.

Households' borrowing constraint and the discount factor wedge. We consider the same environment of the previous example but change the specification of the households' borrowing constraint. The debt limits are now given by $b\left(s^t\right) \geq -\phi_t(z)Y_t(z)$. We assume that z can take two values, $z \in \{z_L, z_H\}$. If $z = z_L$, then debt limits are tight forever as in the previous example, $\phi_t(z) = 0$ for all $t \geq 0$. If $z = z_H$, however, then debt limits are tight at $t = 0$, but agents can borrow a fraction $\phi > 0$ of aggregate output from period $t = 1$ onward. Thus, the realization of z at date $t = 0$ determines to what extent households can hedge *future* income shocks by borrowing. We simplify the idiosyncratic income process and assume that e_t can take two values with equal probability, e_H and e_L with $e_H > e_L$.

In period 0, because $\phi_0(z) = 0$ irrespective of the realization of z, all households are hand-to-mouth and consumption is proportional to the idiosyncratic shock, $c_0(e_0, z) = e_0 C_0(z)$. The same is true from period 1 onward if $z = z_L$. Under the more relaxed debt limit, instead, households that have a negative income shocks can borrow from households with positive income shocks, so the allocation will be different. In period 1, for example, we have $c(e_0, e_H, z_H) = [e_H - \Delta] C_1(z_H)$ and $c(e_0, e_L, z_H) = [e_L + \Delta] C_1(z_H)$ for $\Delta > 0$.[12] Thus, households' consumption at date $t = 1$ is less volatile when $z = z_H$ than when $z = z_L$.

As in the previous analysis, households' precautionary motives depend on the realization of the aggregate shock. If $z = z_L$, households face more volatile consumption in the future and have higher precautionary savings motives, which puts downward pressures on interest rates and output. If $z = z_H$, households have a lower precautionary savings motives, so output and interest rates are higher.

The RA representation captures these effects via the discount factor wedge, which is higher when $z = z_L$ than when $z = z_H$. To see why, consider the households that at date zero have high income. These are those with the highest incentives to save and they are thus the ones achieving the maximum in equation (21). Using their consumption, we can

[12] In equilibrium, $\Delta = b(e_0, e_H, z_H)/C_1(z_H) > 0$ if $\phi > 0$.

express the discount factor wedge as

$$\beta\left(e_H, z_L\right) = \frac{1}{2}\frac{e_H}{e_H} + \frac{1}{2}\frac{e_H}{e_L} > \frac{1}{2}\frac{e_H}{e_H - \Delta} + \frac{1}{2}\frac{e_H}{e_L + \Delta} = \beta\left(e_H, z_H\right).$$

So, higher precautionary motives induced by a tightening of future debt limits are isomorphic to a higher discount factor wedge in the RA representation.[13]

3.4 The role of hand-to-mouth consumers

A series of recent papers has emphasized the critical role of agents with high marginal propensity to consume (MPCs) for the amplification of aggregate shocks in New Keynesian models with incomplete markets, see for example Auclert, Rognlie, and Straub (2018) and Kaplan and Violante (2022). At first sight, it appears that these considerations are not factored in the RA representation: after all, the discount factor wedge is computed using only the consumption share of unconstrained households—those with low MPCs. As we show in this subsection, however, this is not true, as the discount factor wedge can depend on the behavior of constrained households because of general equilibrium.

To explore this point, we consider an economy with MPCs heterogeneity. In this economy, a transfer from low to high MPCs households increases output, an effect that is stronger the larger is the share of high MPCs households in the population. We then derive the RA representation and show two results. First, the discount factor wedge declines with the transfers. Second, the *sensitivity* of the discount factor wedge to the transfer increases in the share of high MPCs households.

The economy last two periods, $t = 0, 1$, and there are only two levels for idiosyncratic productivity in period 0, $e_0 \in \{e_L, e_H\}$. We let μ_L be the probability of drawing e_L. In period 1, all agents have $e_1 = 1$. As in the previous examples, we assume that households can trade only the nominal bond. The household's budget constraints are,

$$c_{i,0} + b_i \leq e_i w_0 l_{i,0} + T_0$$

$$c_{i,1} \leq w_1 l_{i,1} + b_{i,1}(1 + r) - T_{i,1}.$$

We assume that households face tight debt limits, $b_i \geq 0$ and prices are assumed to be fully sticky in period 0 and perfectly flexible in period 1, $\kappa_0 = \infty$ and $\kappa_1 = 0$. To simplify further, the monetary authority sets nominal interest rates so that $\beta\left(1 + i\right)\frac{P_0}{P_1} = 1$, implying a constant real rate equal to $1/\beta$.

[13] The labor supply wedge is equal to one at time zero—as in the previous example— because households are hand-to-mouth and $\sigma = 1$.

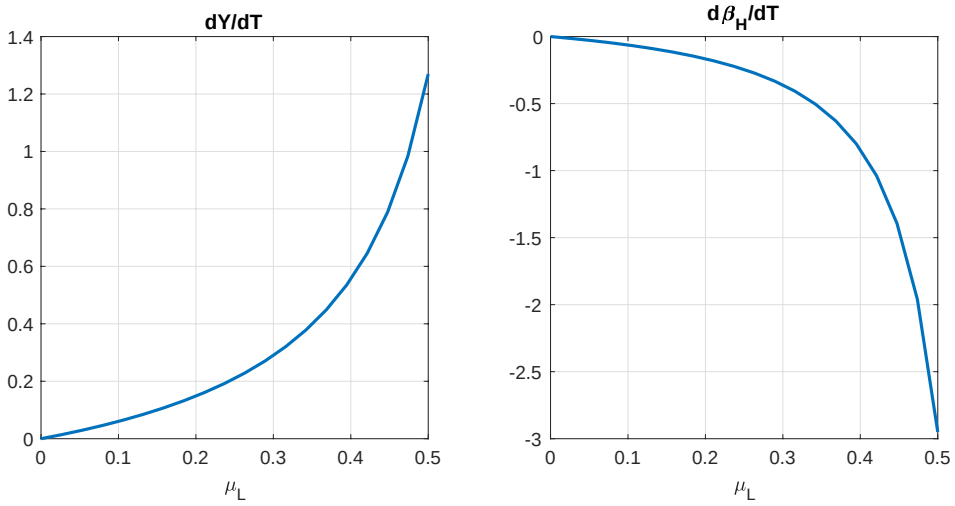


Figure 1: Sensitivity of aggregate consumption and discount factor wedge to transfers

Fiscal policy works as follows. In period 0 the government taxes firms' profits and issues debt B, redistributing the proceed in a lump-sum fashion to all households. In period 1 the government taxes firms' profits and households to repay the debt issued in period 0. The taxes are lump-sum and type-specific. The government budget constraints are

$$T_0 = B + \Pi_0$$

$$\sum_i \mu_i T_{i,1} = B(1+r) + \Pi_1.$$

Market clearing requires that $B = \sum_i \mu_i b_i$. To simplify the algebra, we further assume that the type specific taxes in period 1 are $T_{i,1} = (1+r)b_i + \Pi_1$, so agents have the same consumption in the last period. We assume $\sigma = \psi = 1$ and choose χ so that $c_{i,1} = 1$ for all i. From this it follows that $C_1 = 1$.

A full characterization of this example is presented in Online Appendix B. Because $e_L < e_H$, the debt limit binds for low productivity households if T_0 is small enough. These households are effectively hand-to-mouth, as they consume all of the additional transfer they may receive. In contrast, the consumption of high productivity households is constant over time because $(1+r) = 1/\beta$, and equal to 1 because $c_{H,1} = 1$. Thus, since consumption of high productivity households does not move with transfers while the consumption of low productivity households moves one for one with transfers, the higher is the share of agents with high marginal propensity to consume, μ_L, the larger is the response of aggregate consumption and output to T_0. This is illustrated in the left panel of Figure 1, where we set $\beta = 0.9$ for illustration.

Let us now consider how the RA representation captures these effects. The aggregate

Euler equation is

$$\frac{1}{C_0} = \beta\hat{\beta}_H\,(1+r)\,\frac{1}{C_1}, \tag{27}$$

where $\hat{\beta}_H$ is the discount factor wedge for the unconstrained households—those with high productivity in period 0,

$$\hat{\beta}_H = \frac{C_1}{c_{H,1}} \Big/ \frac{C_0}{c_{H,0}} = \frac{c_{H,0}}{C_0} = \frac{c_{H,0}}{(1-\mu_L)\,c_{H,0} + \mu_L c_{L,0}} = \frac{1}{(1-\mu_L) + \mu_L c_{L,0}}.$$

From the above expression we can see that the discount factor wedge depends on the consumption of constrained households and on their share in the population, even though we use only the consumption share of unconstrained households to compute it. We can then use the RA representation to study the aggregate effects of T_0. The higher the transfer, the higher the consumption of constrained households, the lower is the discount factor wedge $\hat{\beta}_H$. In addition, the higher the share of high MPCs households, μ_L, the larger the sensitivity of the discount factor wedge to T_0, a result that is illustrated in the right panel of Figure 1. The falls in $\hat{\beta}_H$ then leads to an increase in aggregate consumption via (27).[14]

This example illustrates more generally how the RA representation captures the amplification mechanism that takes place in the canonical "two-agent" New Keynesian model.[15] As explained in Bilbiie (2020), this model produces amplification when the income of hand-to-mouth households responds more than one-for-one to a change in aggregate income. So, when aggregate income falls because of some shock, the consumption of hand-to-mouth households falls by more, and this sets in motion the amplification mechanism. The RA representation captures the amplification via the discount factor wedge: the fact that the consumption of hand-to-mouth falls by more than aggregate income implies that the consumption *share* of the unconstrained households increases, which result in an increase in the discount factor wedge.

4 Measuring the wedges

As we discussed in Section 3.2, we need a procedure to measure the wedges and to approximate their stochastic process in order to use the RA representation for counterfactual analysis. We now explain how we accomplish these two steps. In Section 4.1 we show

[14] In this example, $\beta(1+r) = 1$ and $C_1 = 1$. So, the aggregate Euler equation (27) can also be written as $1/C_0 = \hat{\beta}_H$, describing an inverse relation between C_0 and $\hat{\beta}$.

[15] This model features limited participation in the bond market, so it is not formally nested in our framework. However, it still has an RA representation in line with that of Proposition 1. See Online Appendix A.5 for this derivation.

how we can use panel data to estimate the realization of the wedges, while Section 4.2 introduces a class of stochastic processes that we use to approximate their law of motion.

4.1 Measuring the wedges from panel data

Let us denote by i a household with an history of idiosyncratic shocks v^t. We assume that we observe a panel of N households' consumption choices and wage per worked hours, $\{c_{i,t}, w_{i,t}\}$ and the time series for aggregate consumption and wages, $\{C_t, W_t\}$. We also assume that we observe households' financial assets and liabilities.[16] In what follows, we show how we can use these observations to recover the realization of the wedges.

Let us start with the discount factor wedge. From equation (19) we have that the discount factor wedge is the conditional expectation, across realizations of the idiosyncratic state, of the change in the consumption share between two periods raised to a power of $-\sigma$. If we observed multiple histories of consumption choices for the same households we could compute this statistic for each household i by averaging across these different consumption histories. This approach is clearly not feasible because we observe only one consumption path for each household, so we need an alternative way of estimating this conditional expectation.

We proceed as follows. For each household in the panel we compute the consumption shares $\varphi_{i,t} = c_{i,t}/C_t$. We then group households with similar characteristics at date t, leaving us with G groups. For each of group $g \in G$, we compute the statistic

$$\beta_{g,t+1} \equiv \frac{1}{N_g} \sum_{i=1}^{N_g} \left(\frac{\varphi_{i,t+1}}{\varphi_{i,t}} \right)^{-\sigma}, \tag{28}$$

where N_g is the number of households in group g at time t.

The logic behind this approach builds on two premises. The first is that by grouping households along observable characteristics we are proxying for the individual history v^t. The second is that the size of the groups is large enough so that $\beta_{g,t+1}$ approximates the conditional expectation in equation (19) by the law of large numbers.

In our application of Section 5, we will consider partitions based on households' labor income and net worth: at date t we group households according to whether their labor income is above or below median income and, within each of these two groups, whether their net worth is above or below the group median. Thus, for each t, we end up with four different groups of households of approximately equal size: *low income/low net worth*,

[16] In our application these data series will be residualized in order to control for demographic factors and other households' characteristics that are not included in our model.

low income/high net worth, *high income/low net worth* and *high income/high net worth*. For each group g, we use equation (28) to construct $\beta_{g,t+1}$. The rationale behind this partition is that income and net worth are sufficient statistics for an individual history v^t in benchmark incomplete market economies.

Let us now turn to the measurement of ω_t. From equation (20) we can see that we need two households-level observations to construct this series: the consumption shares $\varphi_{i,t}$ and the idiosyncratic productivity $e_{i,t}$. In the class of models described in Section 2, idiosyncratic productivity equals the ratio between the wage per hour of household i and the average hourly wage in the economy, $e_{i,t} = w_{i,t}/W_t$. Given $\{\varphi_{i,t}, e_{i,t}\}$, we can then compute the wedge ω_t using the expression

$$\omega_t = \left[\frac{1}{N}\sum_{i=1}^{N} \varphi_{i,t}^{-\frac{\sigma}{\psi}} e_{i,t}^{\frac{1+\psi}{\psi}}\right]^{-\frac{1}{\psi}}. \tag{29}$$

If N is large enough, this expression is equivalent to the one in (20).

Finally, consider ω_t^{cm} defined in equation (26). To compute this object, we need the observed idiosyncratic productivity and the counterfactual behavior for the consumption shares in the economy with complete markets. Given our assumptions on households' preferences, consumption shares are not time-varying when markets are complete, but they can be potentially different across households because of initial heterogeneity in wealth. Therefore, to compute ω_t^{cm}, we need to know the initial distribution of consumption shares and its correlation with $e_{i,t}$ for every t. In our application we will assume that the moments of the initial distribution are those of the first year in our sample. That is, we compute ω_t^{cm} as follows

$$\omega_t^{\text{cm}} = \left[\frac{1}{N}\sum_{i=1}^{N} \varphi_{i,1}^{-\frac{\sigma}{\psi}} \times \frac{1}{N}\sum_{i=1}^{N} e_{i,t}^{\frac{1+\psi}{\psi}} + \text{Cov}\left(\varphi_{i,1}^{-\frac{\sigma}{\psi}}, e_{i,1}^{\frac{1+\psi}{\psi}}\right)\right]^{-\frac{1}{\psi}}. \tag{30}$$

4.2 Stochastic process for the wedges

As in Chari, Kehoe, and McGrattan (2007), we assume a Markov structure for the states, $\Pr(s_t|s^{t-1}) = \Pr(s_t|s_{t-1})$, and suppose that the equilibrium outcome is induced by a recursive competitive equilibrium. Under these assumptions, $\{\beta_{i,t+1}, \omega_{t+1}, \omega_{t+1}^{\text{cm}}\}$ are functions of the aggregate state variables of the model. In what follows, we specify a class of stochastic process for the wedges based on a first-order approximation of these functions.

In a recursive competitive equilibrium, endogenous variables are functions of idiosyncratic and aggregate states. Let (z, X) be the current realization of the aggregate exogenous and endogenous states, with transition $X' = \Gamma(X, z)$, and let (v, x) be the exogenous and

endogenous idiosyncratic states of the model, with transition $x' = \gamma(x, v, z, X)$.

To make things more concrete, consider a simple economy nested in the class of models of Section 2. Suppose that households can only save and borrow in the risk-free nominal bond and face a borrowing limit $b_{i,t+1} \geq -\phi$, and assume that their idiosyncratic productivity e is an AR(1) process. In a recursive competitive equilibrium, the exogenous aggregate state is $z = [\theta, A, \varepsilon_m]$, the endogenous aggregate state is $X = [\Psi(e, b), i]$—with Ψ being the distribution of individual productivity and bond holdings, and i the lagged nominal interest rate—and the idiosyncratic state variables would be $v = e$ and $x = b$.

Due to the recursive structure, it is straightforward to derive the implied stochastic processes for the discount factor wedge. Denoting by $\varphi(v, x, z, X)$ the consumption share of an household i with individual characteristics (v, x) in the aggregate state (z, X), we can write the discount factor wedge for this household i as

$$\begin{aligned} \beta_{i,t+1} &= \sum_{v_{t+1}} \Pr(v_{t+1}|v_t, z_{t+1}) \left(\frac{\varphi(v_{t+1}, \gamma(v_t, x_t, z_t, X_t), z_{t+1}, \Gamma(z_t, X_t))}{\varphi(v_t, x_t, z_t, X_t)} \right)^{-\sigma} \\ &= f_{\beta_i}(X_t, z_t, z_{t+1}). \end{aligned}$$

Similarly, we can see that in a recursive competitive equilibrium $\omega_{t+1} = f_\omega(X_t, z_t, z_{t+1})$ and $\omega^{\text{cm}}_{t+1} = f_{\omega^{\text{cm}}}(X_t, z_t, z_{t+1})$.

Letting $\hat{y}_t$ be the log-deviation of variable y_t from its steady state, we can express the law of motion of the wedges $\mathbf{T}_{t+1} = [\hat{\beta}_{1,t+1}, \hat{\beta}_{2,t+1}, \ldots, \hat{\omega}_{t+1}, \hat{\omega}^{\text{cm}}_{t+1}]'$, up to a first-order approximation, as

$$\mathbf{T}_{t+1} = \mathbf{A} \times \hat{X}_t + \mathbf{B} \times \hat{z}_t + \mathbf{C} \times \hat{z}_{t+1}, \tag{31}$$

where the matrices $[\mathbf{A}, \mathbf{B}, \mathbf{C}]$ are functions of the primitives of the model.

While (31) is effectively a first order approximation to the true law of motion for $\mathbf{T}_{t+1}$, there is a practical hurdle in using it in our application: this is because certain elements of z_t and X_t may not be defined in the RA representation. To explain this issue, let us go back to the example discussed earlier. In that heterogeneous agent economy, the distribution $\Psi(e, b)$ is a state variable; however, it does not directly appear in the RA representation.[17] In view of this issue, let us partition z_t and X_t as follows: $z_t = [z^{RA}_t, z^{HA}_t]$ and $X_t = [X^{RA}_t, X^{HA}_t]$, where (z^{RA}_t, X^{RA}_t) denote the aggregate states that are also defined in the equivalent RA economy while (z^{HA}_t, X^{HA}_t) are aggregates states in the heterogeneous agents economy that do not directly appear in the RA representation. We then approximate

[17] That is, in the RA representation the distribution $\Psi(e, b)$ affects aggregate variables through its effect on the wedges.

(31) with

$$\mathbf{T}_{t+1} = \mathbf{\Phi}(\mathbf{L}) \times \mathbf{T}_t + \mathbf{A} \times \hat{X}_t^{RA} + \mathbf{B} \times \hat{z}_t^{RA} + \mathbf{C} \times \hat{z}_{t+1}^{RA} + \varepsilon_{t+1}, \tag{32}$$

where $\mathbf{\Phi}(\mathbf{L})$ is a polynomial in the lag operator and ε_{t+1} are innovations with variance-covariance matrix $\mathbf{\Sigma}$. Essentially, our approach consists in proxying for the missing state variables in (31) with lagged values of $\mathbf{T}_{t+1}$.[18]

4.3 Monte Carlo analysis

In Section 4.1 and 4.2 we have discussed how to measure the wedges using panel data and we have proposed a law of motion to capture their behavior. We have made two main approximations in this process. First, the discount factor wedge is computed by averaging the changes in consumption shares of different households with similar characteristics rather than using different histories of the same individual. Second, we cannot condition on all the relevant state variables when considering the law of motion for the wedges because some of these state variables may not be defined in the RA representation. In Online Appendix C we study whether these approximations work well in practice by performing a Monte Carlo analysis on data simulated from benchmark incomplete markets economies.

We consider two economies, the Krusell and Smith (1998) economy and the Guerrieri and Lorenzoni (2017) economy. In both cases, we solve the model and simulate 500 panel datasets comparable to the one we will use in our application to the U.S. economy—10000 households for 25 years. For each of these panel datasets, we proceed in three steps. First, we use households' observations to construct the wedges following the same approach described in Section 4.1. Second, we use the realization of the wedges to estimate the stochastic process in equation (32) with just one lag in $\mathbf{\Phi}(\mathbf{L})$. Third, given the estimated stochastic process for the wedges, we solve for the behavior of aggregate variables using the RA representation. Our test consists then in comparing moments for aggregate variables computed using the RA representation with the actual moments from the heterogeneous agent economy.

Online Appendix C describes in details these steps and provide the results of these experiments. There, we show that our approach to measure the wedges and to approximate their stochastic process works extremely well for both economies, and that the moments computed using the RA representation are identical to the ones of the true underlying economy with heterogeneous agents.

[18] These approximations are common in macroeconomics. For example, they are used when deriving the Vector Autoregressive representation of Dynamic Stochastic General Equilibrium models, see Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007).

5 An application to the U.S. economy

We now apply our framework to U.S. data. In Section 5.1 we use the Consumer Expenditure Survey (CE) to measure the wedges. In Section 5.2 we jointly estimate their stochastic process and the structural parameters of the RA representation. Section 5.3 uses the estimated model to asses the impact of imperfect risk sharing for the U.S. business cycle, while Section 5.4 performs an event study of the Great Recession. Section 5.5 concludes by discussing different economic mechanisms that can explain the fluctuations in the wedges that we document in the data.

5.1 Measuring the wedges

We use the CE to collect information on income, expenditures, employment outcomes, wealth and demographic characteristics for U.S. households between 1992 and 2017. Households in the CE report information on consumption expenditures for a maximum of four consecutive quarters, income and employment information are collected in the first and last interview, and wealth information in the last interview only.[19] Online Appendix D provides details on variable definitions and sample selection, presents summary statistics of the underlying micro data, and compares them to previous studies.

The model of Section 2 abstracts from important determinants of consumption and income, for example demographics. In order to improve the mapping between model and data, we use panel regressions to partial out the effects of these possible confounders. Denoting by $\hat{c}_{i,t}$ the log of consumption expenditures for household i, we estimate the following linear equation

$$\hat{c}_{i,t} = \alpha + \gamma' X_i + e_{i,t},$$

where X_i includes dummies for the sex, race, education, age of the head of household and the state of residence. Our measure of residualized consumption is then $c_{i,t} = \exp\{\alpha + e_{i,t}\}$. We repeat this procedure for all variables used in the analysis. All monetary variables are converted in 2000 dollars using the CPI-U and are reported in per-capita terms.

One concern with our analysis is that measurement errors may affect the computation of the wedges. We perform five steps to mitigate this concern. First, we winsorize the variables used in the construction of the wedges at the top and bottom 1% (year-by-year) in order to correct for reporting mistakes that could result in extreme outliers. Second, we follow Vissing-Jørgensen (2002) and use semi-annual changes when computing $c_{i,t+1}/c_{i,t}$

[19] The CE asks questions about how assets and liabilities have changed in the preceding year, which allows us to back-date wealth information. See https://www.bls.gov/opub/mlr/2012/05/art3full.pdf.

in order to minimize time aggregation and category switching concerns. Third, we obtain the change in households' consumption shares, $\varphi_{i,t+1}/\varphi_{i,t}$, by scaling $c_{i,t+1}/c_{i,t+1}$ with corresponding semi-annual changes in aggregate consumption C_{t+1}/C_t from U.S. National Accounts—arguably less subject to measurement errors.[20] Fourth, we de-mean the wedges, a step that helps dealing with the presence of classical measurement errors.[21] Fifth, we explicitly model a measurement error for the wedges when estimating their stochastic process.

In order to measure the wedges, we set the coefficient of relative risk aversion and the Frisch elasticity of labor supply to one, $\sigma = \psi = 1$, conventional values in the literature.

The discount factor wedge. We measure $\beta_{g,t}$ following the approach described in Section 4.1. Each of the four groups of households contains approximately 875 observations per year. For each household in group g we compute the semi-annual change in consumption expenditures following Vissing-Jørgensen (2002) and scale it by the semi-annual change in aggregate consumption over the same horizon. We then square the resulting ratio and obtain an empirical analog to the yearly change in consumption shares, $\varphi_{i,t+1}/\varphi_{i,t}$. The discount factor wedge for group g is then constructed by averaging $(\varphi_{i,t+1}/\varphi_{i,t})^{-1}$ across the households in that group, see equation (28).

Table 1 reports summary statistics of the discount factor wedge for each of the four groups. We report the wedges in log deviations from the sample average of the third group—the high income/low net worth households. There are two features of the data that we wish to emphasize. First, high-income households have on average a higher discount factor wedge than low-income households. Focusing on households with low net-worth, the discount factor wedge of the high income group is 300 basis points higher than the discount factor wedge of the the low income group. Second, the discount factor wedge for the high income/low net worth group is countercyclical, with a correlation coefficient of -0.15 with U.S. detrended output, while it is procyclical for the other groups.

The fact that high income/low net worth households have a relatively high and countercyclical discount factor wedge makes it a natural choice to be a group of financially unconstrained households—standing in for households that attain the maximum in equa-

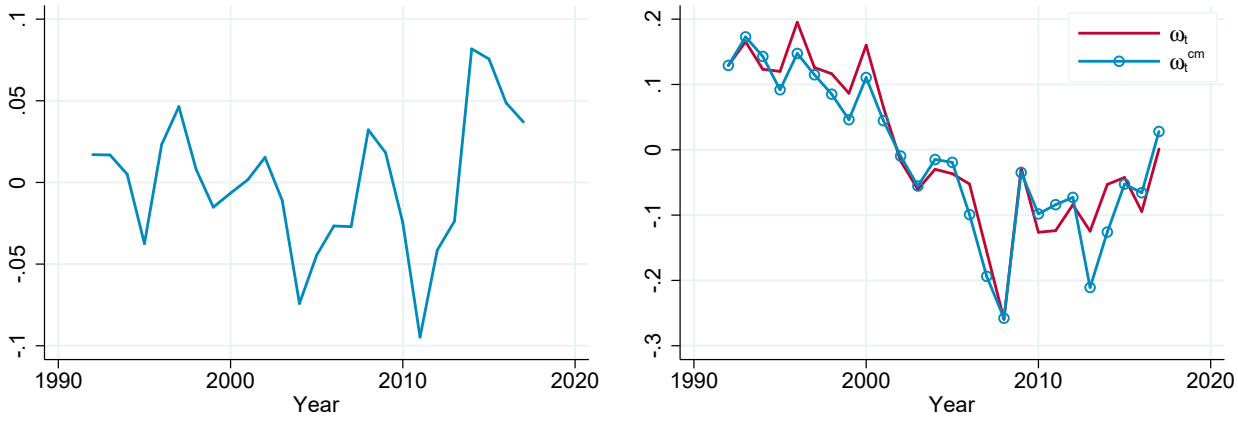
[20] Our results are comparable when we instead compute aggregate consumption using the cross-sectional average of consumption expenditures in the CE, see Online Appendix D.5.

[21] To understand why de-meaning helps, consider the construction of the discount factor wedge. Suppose that observed consumption is related to the true consumption as follows, $c_{i,t} = c_{i,t}^{\text{true}} \times \exp\{\eta_{i,t}\}$ where $\eta_{i,t} \sim \mathcal{N}(-\sigma_\eta^2/2, \sigma_\eta^2)$ is a classical measurement error. Simple calculations, then, show that $\beta_{g,t}$ defined in (28) converges to $\beta_{g,t}^{\text{true}} \times \exp\{\sigma_\eta^2\}$ when N_g is large. So, de-meaning $\beta_{g,t}$ (in logs) removes the bias due to measurement error. A similar derivation can be done for the labor disutility wedge.

Table 1: Summary statistics for the discount factor wedge

	$\text{Mean}(\hat{\beta}_{g,t})$	$\text{Corr}(\hat{\beta}_{g,t}, \hat{\beta}_{g,t-1})$	$\text{Stdev}(\hat{\beta}_{g,t})$	$\text{Corr}(\hat{\beta}_{g,t}, \hat{Y}_t)$
Low income/low net worth	-0.03	0.38	0.04	0.12
Low income/high net worth	-0.02	0.63	0.05	0.25
High income/low net worth	0.00	0.53	0.04	-0.15
High income/high net worth	0.01	0.33	0.04	0.12

Notes: *For each t, we partition households into four group depending on their income and net worth at date t as described in Section 4.1. We then compute for each household the inverse change in consumption share between t and $t+1$, $(\varphi_{i,t+1}/\varphi_{i,t})^{-1}$. For each group we then compute $\beta_{g,t+1}$ using equation (28). We scale each $\beta_{g,t+1}$ by the sample average of the high income/low net worth group, and report the normalized wedges in logs. See Online Appendix D for the definition of detrended output $\hat{Y}_t$.*



(a) Discount factor wedge

(b) Labor disutility wedge

Notes: *Authors' calculation based on the the CE. Panel (a) plots $\log(\beta_{g,t})$ defined in equation (28) for the group of households with high-income/low net worth. Panel (b) plots $\log(\omega_t)$ and $\log(\omega_t^{cm})$, defined respectively in equation (29) and (30). The series are normalized to have a sample average of zero.*

Figure 2: The time path of the wedges

tion (21). Indeed, we have that

$$\max_i \mathbb{E}_t \left[\beta_{i,t+1} \frac{C_t}{C_{t+1}} \right] = \max_i \left\{ \mathbb{E}_t[\beta_{i,t+1}] \mathbb{E}_t \left[\frac{C_t}{C_{t-1}} \right] + \text{Cov}_t \left(\beta_{i,t+1}, \frac{C_t}{C_{t+1}} \right) \right\}.$$

So, groups that on average have a high and countercyclical discount factor wedge are likely to be those achieving the maximum in (21). This empirical finding is also in line with standard incomplete market models: households that experience positive income shocks and that have low net worth have the highest incentives to save, and thus are not financially constrained. Because of these reasons, we will use $\beta_{g,t}$ of this group as the discount factor wedge in our analysis.

Panel (a) of Figure 2 plots the time-series of the discount factor wedge. We can see that the discount factor wedge increases substantially during the Great Recession and remains elevated in the post 2010 period.

The labor disutility wedge. We compute for each household in our panel the consumption share $\varphi_{i,t}$ and the ratio between the household's hourly wage and the average hourly wage in the panel, $e_{i,t} = w_{i,t}/W_t$. We then use equations (29) and (30) to obtain the time path of ω_t and ω_t^{cm}.

Figure 2 panel (b) plots these two series. We can see for both series a clear downward trend in the early part of the sample. The downward trend is explained by the increase in the cross-sectional variance of wages for U.S. households during this period, a fact that is well established in the literature (Heathcote, Perri, and Violante, 2010). As we explained in Section 3.2, an increase in the cross-sectional dispersion of labor productivity leads high-productivity workers to increase their labor supply relative to low productivity workers because of substitution effects. This mechanism operates irrespective of whether financial markets are complete or not, and the resulting change in the composition of the labor force that takes place in the heterogeneous agent economy is captured in the RA representation by a decline in the labor disutility wedge. So, both ω_t and ω_t^{cm} fall for large part of the sample. We can also observe from the figure that the deviations between ω_t and ω_t^{cm} are typically small. This means that the wealth effects in labor supply that are due to uninsured idiosyncratic income risk—and that our framework captures with a discrepancy between ω_t and ω_t^{cm}—are quite small.

5.2 Estimating the RA representation

We now estimate the stochastic process of the wedges and the structural parameters of the RA representation. The structural shocks in the RA representation is $z_t^{RA} = [\hat{\theta}_t, \hat{A}_t, \varepsilon_{m,t}]$. We assume that the aggregate preference and technology shocks in logs follow independent AR(1) processes,

$$\begin{aligned} \hat{\theta}_t &= \rho_\theta \hat{\theta}_{t-1} + \varepsilon_{\theta,t} \qquad & \varepsilon_{\theta,t} \sim \mathcal{N}(0, \sigma_\theta^2) \\ \hat{A}_t &= \rho_a \hat{A}_{t-1} + \varepsilon_{A,t} \qquad & \varepsilon_{A,t} \sim \mathcal{N}(0, \sigma_A^2), \end{aligned}$$

and that monetary policy innovations are Gaussian, $\varepsilon_{m,t} \sim \mathcal{N}(0, \sigma_m^2)$. The endogenous state variables are the nominal interest rate and the wedges, $\mathbf{T}_t = [\hat{\beta}_t, \hat{\omega}_t, \hat{\omega}_t^{\text{cm}}]'$. The law of motion for $\mathbf{T}_t$ is given by equation (32). We restrict $\mathbf{\Phi}(L)$ to have a one-lag structure and to be block diagonal, so that ω_t^{cm} depends only on its own lags and does not load on the

other equations. The other parameters in (32) are left unrestricted.

We estimate the parameters of the stochastic processes jointly with the other structural parameters governing preferences, technology and the behavior of the monetary authority. We fix a subset of these parameters to conventional values in the literature. Consistent with the measurement of the wedges, we set $\sigma = \psi = 1$. We let $\mu = 1.2$ and set χ to $1/\mu$, so that consumption and output are normalized to 1 in a deterministic steady state of the model. Finally, we set the target inflation rate to 2%, and $\beta = 0.99$, values that guarantee that the model roughly matches the average inflation and nominal interest rate in our sample.

The remaining model parameters, which we collect in the vector Θ, are estimated with Bayesian methods using annual data on output, inflation, nominal interest rates and the wedges. We map the log of output, $\hat{Y}_t$, to detrended log real Gross Domestic Product, the inflation rate π_t to the annual percent change in the Consumer Price Index, and nominal interest rates i_t to the annual effective Federal Funds Rate, see Online Appendix D for definitions, data sources and de-trending methodology. The sample period is 1992-2017. We denote by $\mathbf{Y}_t = [\hat{Y}_t, i_t, \pi_t, \mathbf{T}_t]$ the observables at time t and by $\mathbf{S}_t = [i_{t-1}, \hat{\theta}_t, \hat{A}_t, \varepsilon_{m,t}, \mathbf{T}_t]$ the state vector. The RA representation of Proposition 1 defines implicitly a law of motion for these vectors,

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{g}(\mathbf{S}_t; \Theta) + \eta_t \\ \mathbf{S}_t &= \mathbf{f}(\mathbf{S}_{t-1}, \varepsilon_t; \Theta), \end{aligned} \tag{33}$$

where $\mathbf{g}(.)$ and $\mathbf{f}(.)$ represent the policy functions of the RA representation, ε_t collects the innovations to the stochastic variables of the model, and η_t are Gaussian measurement errors. We introduce measurement errors only for the wedges and fix their variance to 10% of the unconditional variance of these series. For the purpose of estimation, the policy functions are approximated using a first-order perturbation. When performing the counterfactuals, however, we solve the model with global solution methods that allow for the possibility of a binding zero lower bound constraint on nominal interest rates, see Online Appendix E.[22]

The posterior distribution of Θ is characterized using a Random Walk Metropolis Hastings algorithm, see An and Schorfheide (2007). Online Appendix F provides details on the estimation algorithm and it reports parameters' estimates and indicators of model fit. Tables A-4 and A-5 report prior and posterior statistics for the model parameters. Our estimates for the parameters of the Taylor rule and of the price adjustment costs are in line

[22] We estimate a log-linearized version of the model because the numerical solution is much faster and numerically more stable than the global approximation that we use for the counterfactuals. See Borağan Aruoba, Cuba-Borda, and Schorfheide (2018) for a similar approach.

with previous studies. In addition, we estimate some spillovers between the model state variables and the wedges. Recessionary shocks at time t—such as a negative technology shock or a positive aggregate discount factor shock—forecast an increase in the discount factor wedge at time $t+1$. This result is consistent with heterogeneous agents economies where households' precautionary savings motives are more prevalent in recessions. Finally, the parameters of the stochastic process for $\hat{\omega}_t^{\text{cm}}$ are close to those of $\hat{\omega}_t$, which is not surprising given that the two series display almost an identical pattern over the sample. Figure A-5 reports model implied distributions for the mean, standard deviation, autocorrelation and cross-correlation of output, inflation and nominal interest rates and compare these to the corresponding sample moments. We can see from the figure that the estimated model fits reasonably well the behavior of these series, as sample statistics from the data lie within the corresponding model implied distribution.

5.3 Imperfect risk sharing and the business cycle

In this and the next subsection we will compare the estimated RA representation to an economy that is identical to the RA representation with the exception that $\hat{\beta}_t = 0$ and $\hat{\omega}_t = \hat{\omega}_t^{\text{cm}}$. As we showed in Proposition 2, if we set the wedges to those values in the RA representation we would recover the path of aggregate variables in presence of complete financial markets. So, we will refer to this economy as the *complete markets counterfactual* (CM). By comparing these two economies, we are able to assess the role of imperfect risk sharing for macroeconomic aggregates. The analysis will focus on comparing the business cycle properties of the two economies and not their long run average behavior.[23]

We set the model parameters at the posterior mean, solve numerically for the RA representation and the complete markets counterfactual, and compute long simulations for both economies. Table 2 reports key statistics for output, inflation and nominal interest rates. If shocks and frictions at the micro level were an important source of business cycle fluctuations, we should expect the complete markets counterfactual to display significantly less volatile output than what we find in the RA representation. We can see from Table 2 that the standard deviation of output in CM is approximately 7% smaller relative to that of the RA representation. In addition, by comparing the cross-correlation patterns in the two economies, we can see that imperfect risk sharing act mostly as a "demand" shock on the economy: output, inflation and nominal interest rates are much more positively correlated

[23]Our approach is not designed to study this latter question because, by construction, the complete market counterfactual and the RA representation have the same steady state. This happens because we normalize the wedges to have a mean of 1. In Online Appendix F.4 we show that this restriction does not affect our key results, as the business cycle properties of the RA representation and the CM counterfactual are almost invariant to the steady state value of the wedges.

Table 2: Imperfect risk sharing and the business cycle

	RA	CM	CM, $\omega_t^{\text{cm}} = \omega_t$	RA, $\sigma_\beta = 0$
Stdev$(\hat{Y}_t)$	4.05	3.79	3.82	3.99
Stdev(π_t)	1.39	0.96	0.94	1.36
Stdev(i_t)	2.20	1.90	1.84	2.17
Corr$(\hat{Y}_t, \pi_t)$	0.45	0.12	0.22	0.42
Corr$(\hat{Y}_t, i_t)$	0.17	-0.36	-0.11	0.13
Corr(π_t, i_t)	0.51	0.33	0.31	0.49

Notes: *Each column reports statistics computed from simulations of four different economies: i) the RA representation; ii) the counterfactual economy with complete markets; iii) the counterfactual economy with complete markets but $\omega_t^{cm} = \omega_t$; iv) The RA representation with $\sigma_\beta = 0$. Statistics for each version of the model are computed on a long simulation ($T = 50,000$).*

in the RA representation than in the economy with complete financial markets.

In order to understand these results, it is useful to note that there are two key differences between the RA representation and CM: the latter does not have a discount factor wedge and it has a different labor disutility wedge. In an effort to understand which of these two features drive the results, we report in the third column in Table 2 moments computed from the CM economy but condition the realization of the labor disutility wedge to be the same as in the RA representation, $\omega_t^{\text{cm}} = \omega_t$. By comparing column two and three, we can see that these two versions of the CM economy features almost identical business cycle properties. Therefore, the differences between RA and CM are mostly driven by the movements in the discount factor wedge that are present in the former but not in the latter. This helps explaining why the correlation between output, inflation and nominal interest rates drops in CM, as shocks to the discount factor in the canonical three-equations New Keynesian model induce positive comovement between these variables.

These results suggest that movements in the discount factor wedge are key for understanding the aggregate implications of imperfect risk sharing. An interesting question is whether these fluctuations are mostly induced by the structural shocks, $[A_t, \theta_t, \varepsilon_{m,t}]$, or whether they are due to the innovations $e_{\beta,t}$? The fourth column of Table 2 helps us answering this question. There, we report statistics from the RA economy when the variance of $e_{\beta,t}$ is set to zero, $\sigma_\beta = 0$, so all the movements in the discount factor wedge are due to spillovers from the structural shocks. We can see that the sample moments in this economy are comparable to those of the benchmark RA representation, implying that most of the aggregate effects of imperfect risk sharing arise because incomplete markets amplify the effects of structural shocks.

Next we explore why imperfect risk sharing does not have sizable effects on output on

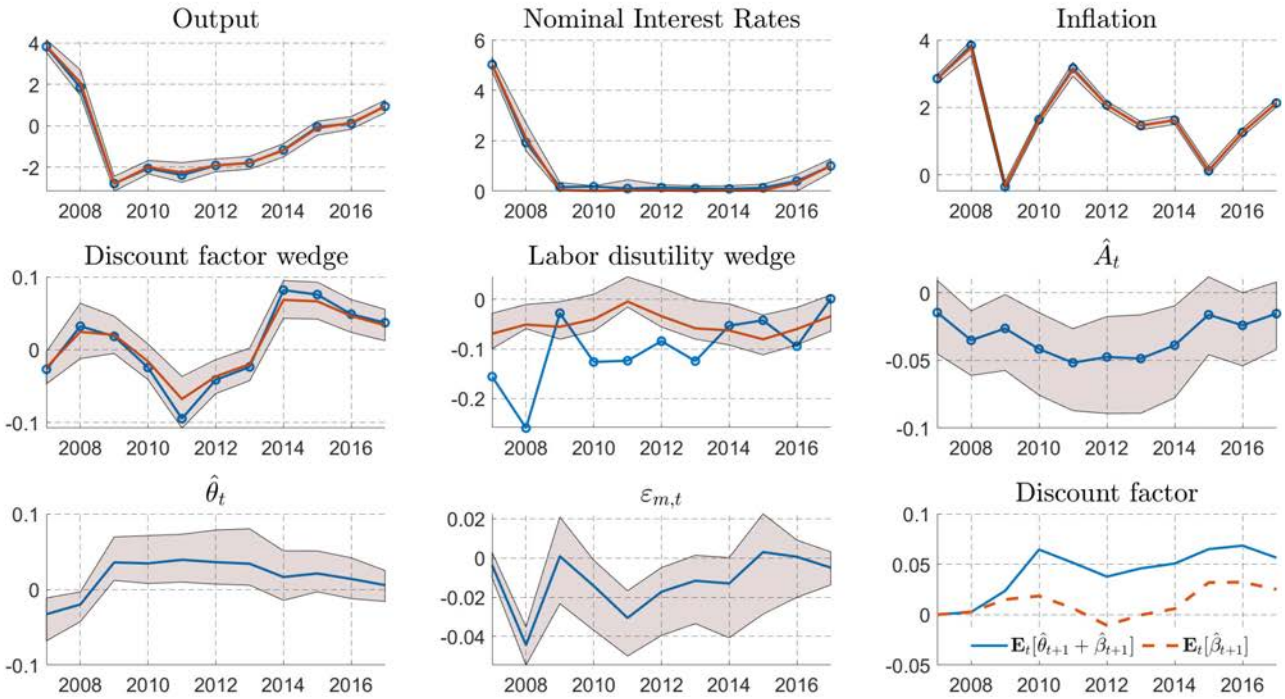
average. The results of this section suggest that the main channel through which micro heterogeneity matters for aggregate variables is equivalent to movements in the discount factor. In the standard New Keynesian model, the aggregate implications of shocks to the discount factor critically depend on the response of the monetary authority. For a fixed nominal interest rate, an increase in the discount factor induces the stand-in household to cut consumption. If the monetary authority responds by lowering nominal interest rates, however, the increase in patience has more limited effects on consumption and output. Given our estimates of the Taylor rule, the response of the monetary authority is strong enough to offset most of the output effects of the measured changes in the discount factor wedge.

Importantly, this result depends on the response of the monetary authority. It is well known in the literature that changes in the discount factor can have substantial output effects when the zero lower bound constraint on nominal interest rate binds, as in that case the monetary authority cannot cut further nominal interest rates. These events are somewhat rare in the estimated model. Thus, while the results of Table 2 indicate limited output effects of imperfect risk sharing *on average* over the business cycle, they do not rule out that these frictions may play a more important role in periods during which the monetary authority is constrained by the zero lower bound. In the next subsection we explore this possibility with an event study of the U.S. Great Recession.

5.4 Imperfect risk sharing and the Great Recession

We now use the estimated model to measure the macroeconomic effects of imperfect risk sharing during the Great Recession. To this end, we proceed in two steps. In the first step, we apply the particle filter to the state-space system (33) and estimate the realization of the structural shocks. In the second step, we feed the structural shocks in the CM economy to obtain the counterfactual paths for output, inflation and nominal interest rates under complete financial markets. The difference between what we observe in the data and these counterfactual paths isolates the macroeconomic effects of imperfect risk sharing during the Great Recession. Online Appendix F provides a detailed description of both steps.

Starting with the first step of this procedure, Figure 3 reports the data (circled lines) along with the posterior mean (solid line) and 90% credible set for their model counterpart. The figure also reports the estimates for the three structural shocks. By construction, the model tracks very closely output, inflation and nominal interest rates during the event. In order to replicate these paths, the model infers a substantial increase in the discount factor of the stand-in household: $\mathbb{E}_t[\hat{\beta}_{t+1} + \hat{\theta}_{t+1}]$ increases by six percentage points during the event. This is a well known result from the literature: the canonical three-equations New



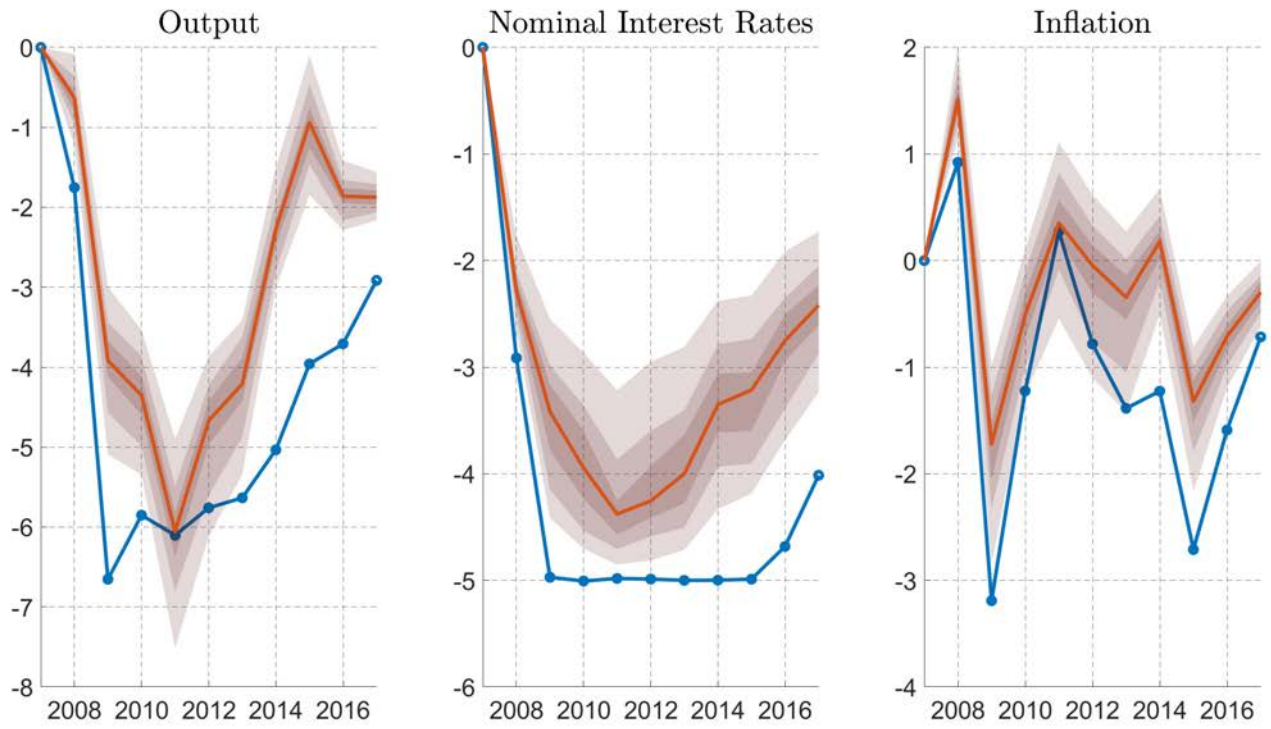
Notes: *The circled line reports the data used in the experiment. The solid lines report the posterior mean of the filtered series for* $\mathbf{Y}_t$ *and* $\mathbf{S}_t$*, while the shaded area reports 90% credible sets (equal tail probability). Inflation and nominal interest rates are reported in percentages, while output is reported in percentage deviations from a quadratic time trend. The remaining variables are in log-deviation from their steady-state value.*

Figure 3: Applying the particle filter to the RA representation

Keynesian model needs an increase in the discount factor in order to fit the fall in output, inflation and nominal interest rate observed after 2008. Interestingly, a substantial fraction of this increase is due to movements in $\mathbb{E}_t[\hat{\beta}_{t+1}]$, roughly three of the size percentage points by the end of the episode.

Equipped with the path for the structural shocks, we can then construct the trajectories for output, nominal interest rates and inflation that would prevail in an economy with complete financial markets. For that purpose, we solve numerically for the policy functions of the CM economy and construct the counterfactual paths by feeding these policy functions with the path for $\{\hat{\theta}_t, \hat{A}_t, \varepsilon_{m,t}\}$ we estimated using the particle filter and with $\hat{\omega}_t^{\text{cm}}$.

Figure 4 compares the trajectories for output, inflation and nominal interest rates in this counterfactual (solid lines) with the actual trajectories in U.S. data (circled lines) during the 2007-2017 period. From 2007 to 2009, output in the US fell by 6.6%. In the counterfactual economy, instead, output falls by only 3.9%. In addition, the figure shows that the recovery from the Great Recession would have been faster with complete financial markets relative to what we have observed in the data. Thus, our analysis supports the view that these financial market frictions were an important driver of the depth and persistence of the U.S.



Notes: *The circled line reports output, nominal interest rates and inflation during the 2007-2017 period. Inflation and nominal interest rates are reported in percentages, while output is reported in percentage deviations from a quadratic time trend. We normalize these variables to 0 in 2007. The solid line reports the posterior mean of the same variable in the counterfactual economy with complete financial markets while the shaded area reports 64% credible sets (equal tail probability).*

Figure 4: Imperfect risk sharing and the U.S. Great Recession

Great Recession.

Why do we observe these differences between the baseline and the counterfactual economy? From Figure 3 we can see that $\mathbb{E}_t[\hat{\beta}_{t+1}]$ contributes substantially to the increase in the overall discount rate during the Great Recession and, because interest rates are at their lower bound from 2009 onward, these developments lead to a substantial decline in aggregate demand. In the CM economy, instead, $\hat{\beta}_t$ is constant and does not increase during this period, implying a higher trajectory for output, inflation and nominal interest rates.

5.5 Inspecting the mechanisms

While our results indicate that shocks and frictions limiting risk sharing across households contributed significantly to the depth and persistence of the Great Recession, so far we have been silent about the specific economic mechanisms responsible for this result. In this subsection we use additional information to understand which economic mechanism can account for our results and, specifically, for the rise in the discount factor wedge.

A first useful exercise consists in understanding which moments of the distribution of households' consumption shares are responsible for the observed movements in the discount factor wedge. From equation (28), we can decompose the log of $\beta_{g,t+1}$ for households' group g as follows

$$\log\left(\beta_{g,t+1}\right) = \underbrace{\Delta\log(C_{t+1}) - \Delta\log(C_{g,t+1})}_{\beta^{\text{Avg}}_{g,t+1}} + \underbrace{\log\left(\frac{1}{N_g}\sum_{i=1}^{N_g}\left[\frac{C_{g,t+1}/C_{g,t}}{c_{i,t+1}/c_{i,t}}\right]\right)}_{\beta^{\text{Jen}}_{g,t+1}}, \tag{34}$$

where $C_{g,t}$ is the average consumption at time t of households in group g. Mechanically, there are two factors that can explain the increase in the discount factor wedge during the Great Recession. First, a persistent fall in the average consumption of households in the high income/low net worth group relative to aggregate consumption, the term $\beta^{\text{Avg}}_{g,t}$ in equation (34). Second, a increase in the cross-sectional dispersion of consumption *within* group g, an effect captured by the "Jensen" term $\beta^{\text{Jen}}_{g,t+1}$. These two effects are intuitive: fixing the interest rate, a household that expects low and/or volatile consumption in the future has more incentives to save today, an effect that is captured in the RA representation by a higher discount factor wedge.

This decomposition is useful for the purpose of model discrimination because two models that generate similar dynamics for $\beta_{g,t}$ may have different implications for the two terms in the right hand side of equation (34).[24] For example, take the "two-agent" New Keynesian model studied in Bilbiie (2008) and Galí, López-Salido, and Vallés (2007). As showed in Debortoli and Galí (2017), these models feature similar macroeconomic dynamics relative to New Keynesian models with richer heterogeneity, such as Kaplan, Moll, and Violante (2018). This means that those models imply a similar behavior for the discount factor wedge $\beta_{g,t}$. However, the two-agent New Keynesian model does not feature idiosyncratic consumption risk for unconstrained households, which implies that $\beta^{\text{Jen}}_t = 1$ and all movements in the discount factor wedge are induced by changes in β^{Avg}_t. Models with richer heterogeneity can instead generate time-variation in β^{Jen}_t.

Figure 5 reports $\log(\beta_{g,t})$ and $\beta^{\text{Jen}}_{g,t}$ for the high income/low net worth group, with $\beta^{\text{Avg}}_{g,t}$ being the difference between these two series. We can see that the dynamics of the discount factor wedge during the Great Recession are driven almost entirely by the Jensen component.[25] Therefore, in order to understand the rise in the discount factor wedge during this

[24] Discriminating between two models is useful even when they imply similar dynamics for the discount factor wedge, as the two models may have very different policy implications.

[25] This is true also for the other group of households that are likely to be financially unconstrained, the high income/high net worth households.

Notes: *This figure plots the decomposition described in equation (34) for high income/low net worth households. The solid line reports* $\log(\beta_{g,t})$ *while the circled line reports* β_t^{Jen}*. The variables are expressed in deviations from their 2007 value.*

Figure 5: Decomposition of equation (34)

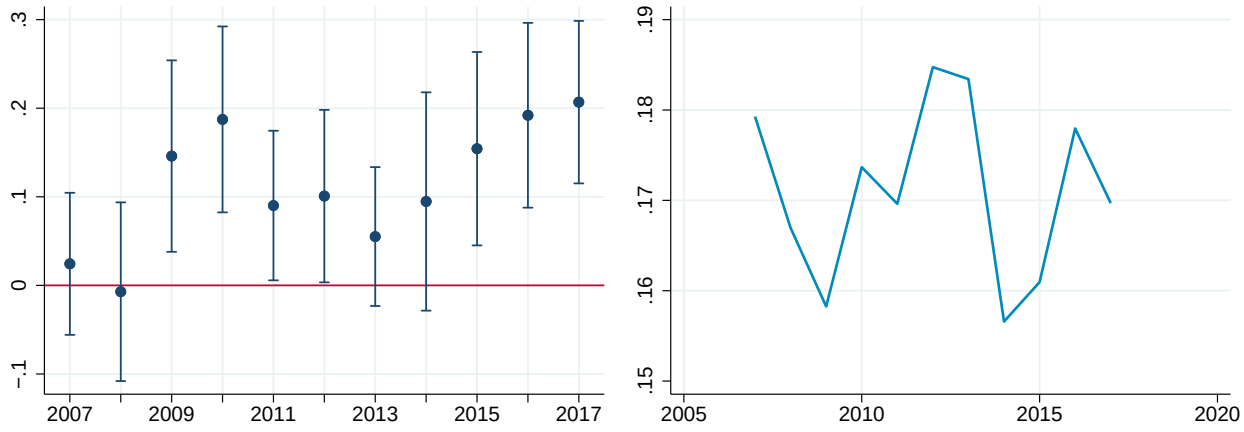
period, we need to focus on economic mechanisms that can generate an increase in the consumption risk of unconstrained households.

The existing literature has emphasized two main mechanisms that can lead to such pattern, which we discussed in the two examples of Section 3.3. In the first example, we showed that an increase in idiosyncratic labor income risk can lead to higher consumption risk and more precautionary savings. In the second example, we showed that similar effects can arise after a persistent tightening of households' borrowing constraints. Both of these forces can rationalize an increase of the "Jensen" term in equation (34) and be consistent with the findings of Figure 5.

In order to distinguish between these two forces, it is useful to consider the following reduced form model of households' consumption growth,

$$\Delta \hat{c}_{i,t} = \tau_t + \lambda_t \Delta \hat{y}_{i,t} + e_{i,t}, \tag{35}$$

where $\hat{y}_{i,t}$ is the log of households' disposable income at date t. In this framework, the conditional variance of consumption growth depends on the volatility of income and on the sensitivity of consumption to income changes, λ_t. It is straightforward to see that the two examples described in Section 3.3 have different implications for these two components. In the first example, we have that $\lambda_t = 1$ for all t; therefore, the volatility of consumption varies over time only because of changes in the volatility of income. In the second example,



(a) Estimates of λ_t (b) Var$(\Delta \hat{y}_{i,t})$

Notes: *For each t, we select households that at $t-1$ are in the group of high income/low net worth households and estimate equation (35) by ordinary least squares. Panel (a) plots the point estimate for λ_t and 90% robust confidence interval. Panel (b) plots the cross-sectional variance of $\Delta \hat{y}_{i,t}$ for each t.*

Figure 6: Drivers of households' consumption volatility, 2007-2017

instead, income volatility is constant over time while shocks to households' borrowing constraint affect the sensitivity of their consumption to income shocks; therefore, consumption volatility changes over time only because of movements in λ_t.

We estimate equation (35) for the high income/low net worth households and report the results in Figure 6. Panel (a) plots the estimated coefficient λ_t over time while panel (b) plots the variance of $\Delta \hat{y}_{i,t}$ for this group of households. From the figure we can see that λ_t tracks well the dynamics of $\beta_{g,t}^{\text{Jen}}$ reported in Figure 5. The cross-sectional variance of income for these households, instead, does not display a similar pattern and it is fairly flat over this horizon.[26]

Overall, these results suggest that the increase in the discount factor wedge during the Great Recession is mostly the result of an increase in consumption volatility of unconstrained households, and that this was mostly driven by an increase in the sensitivity of their consumption to income changes. When interpreted through the lens of the examples of Section 3.3, this evidence favors economic models that emphasize a disruption of

[26]This result is consistent with previous research that examined the cyclical behavior of income changes. While Storesletten, Telmer, and Yaron (2004) found evidence of countercyclical income volatility, their findings are not comparable to those of Figure 6 because we are conditioning on high income/low net worth households while Storesletten, Telmer, and Yaron (2004) use their entire sample in their analysis. Consistent with our findings, Heathcote, Perri, and Violante (2010) find little cyclical variation in the volatility of earnings growth for households between the 50th and 90th percentiles of the earnings distribution. These authors document significant cyclical variation in income volatility for poorer households due to a higher incidence of unemployment during recessions, a result confirmed by Guvenen, Ozkan, and Song (2014).

households' ability to insure fluctuations in their income during the Great Recession, as in Guerrieri and Lorenzoni (2017). It is worth noting, however, that this interpretation is somewhat special to these examples. In general, changes in the income process may affect the sensitivity of consumption to income—λ_t in equation (35)—and potentially account for the dynamics reported in Figure 6. For example, suppose that the Great Recession was accompanied by an increase in the variance of the permanent component of income and a reduction of its transitory component.[27] Such a shift could explain why in Figure 6 we detect an increase in λ_t in absence of significant movements in Var $(\Delta\tilde{y}_{i,t})$. This is because permanent shocks are harder to smooth than transitory ones in baseline incomplete markets model, and so consumption is more sensitive to the former than the latter.

6 Conclusion

We have developed an approach to assess the macroeconomic implications of imperfect risk sharing implied by a class of New Keneysian models with heterogeneous agents. In this class of models, households' inability to insure idiosyncratic risk implies time-variation in their consumption shares. Leveraging this insight, we use households' consumption choices to directly measure the degree of imperfect risk sharing for the U.S. economy. We have documented a deterioration of risk sharing during the Great Recession, as the cross-sectional dispersion of households' consumption shares increases during this period. We have then proposed a methodology to quantify the aggregate implications of these movements. Through the lens of a prototypical New Keynesian model, these deviations from perfect risk sharing contribute little to the business cycle on average, but explains a substantial part of the depth and persistence of the Great Recession.

Our paper clarifies that assumptions about the nature of the idiosyncratic risk faced by households and about the private and public risk sharing mechanisms available to them matter for aggregate fluctuations only through their impact on two wedges. These wedges depend on the joint distribution of households' consumption shares and relative wages, and they can be computed using panel data. We believe they are useful empirical targets for the analysis of heterogeneous agents economies.

[27] This hypothesis is similar to the mechanism in Blundell, Pistaferri, and Preston (2008) to account for the changes in the consumption distribution relative to the income distribution over the 1970s to 1990s.

References

Acharya, Sushant and Keshav Dogra. 2018. "Understanding HANK: Insights from a PRANK." *Federal Reserve Bank of New York Staff Report* 835.

Alvarez, Fernando and Urban J Jermann. 2000. "Efficiency, equilibrium, and asset pricing with risk of default." *Econometrica* 68 (4):775–797.

An, Sungbae and Frank Schorfheide. 2007. "Bayesian analysis of DSGE models." *Econometric reviews* 26 (2-4):113–172.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New trade models, same old gains?" *American Economic Review* 102 (1):94–130.

Aruoba, S Borağan, Luigi Bocola, and Frank Schorfheide. 2017. "Assessing DSGE model nonlinearities." *Journal of Economic Dynamics and Control* 83:34–54.

Auclert, Adrien. 2017. "Monetary policy and the redistribution channel." Tech. rep., National Bureau of Economic Research.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. 2018. "The Intertemporal Keynesian Cross." Tech. rep., January 2017 working paper.

———. 2020. "Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model." Tech. rep., National Bureau of Economic Research.

Barro, Robert J and Robert G King. 1984. "Time-separable preferences and intertemporal-substitution models of business cycles." *The Quarterly Journal of Economics* 99 (4):817–839.

Bayer, Christian, Benjamin Born, and Ralph Luetticke. 2020. "Shocks, frictions, and inequality in US business cycles." Tech. rep., CEPR Discussion Paper No. DP14364.

Bayer, Christian, Ralph Lütticke, Lien Pham-Dao, and Volker Tjaden. 2019. "Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk." *Econometrica* 87 (1):255–290.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J Sargent. 2018. "Inequality, business cycles, and monetary-fiscal policy." Tech. rep., National Bureau of Economic Research.

Bilbiie, Florin, Giorgio E Primiceri, and Andrea Tambalotti. 2022. "Inequality and Business Cycles." Tech. rep., Unpublished manuscript.

Bilbiie, Florin O. 2008. "Limited asset markets participation, monetary policy and (inverted) aggregate demand logic." *Journal of Economic Theory* 140 (1):162–196.

———. 2020. "The new keynesian cross." *Journal of Monetary Economics* 114:90–108.

Bilbiie, Florin Ovidiu. 2021. "Monetary policy and heterogeneity: An analytical framework." .

Bisin, Alberto, Gian Luca Clementi, and Piero Gottardi. 2017. "Capital Structure and Hedging Demand with Incomplete Markets." .

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2008. "Consumption inequality and partial insurance." *American Economic Review* 98 (5):1887–1921.

Boppart, Timo, Per Krusell, and Kurt Mitman. 2018. "Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative." *Journal of Economic Dynamics and Control* 89:68–92.

Borağan Aruoba, S, Pablo Cuba-Borda, and Frank Schorfheide. 2018. "Macroeconomic dynamics near the ZLB: A tale of two countries." *The Review of Economic Studies* 85 (1):87–118.

Brav, Alon, George M Constantinides, and Christopher C Geczy. 2002. "Asset pricing with heterogeneous consumers and limited participation: Empirical evidence." *Journal of Political Economy* 110 (4):793–824.

Broer, Tobias, Niels-Jakob H. Hansen, Per Krusell, and Erik Oberg. 2018. "The New Keynesian Transmission Mechanism: A Heterogeneous-Agent Perspective." *Review of Economic Studies* forthcoming.

Buera, Francisco J and Benjamin Moll. 2015. "Aggregate Implications of a Credit Crunch: The Importance of Heterogeneity." *American Economic Journal: Macroeconomics* 7 (3):1–42.

Challe, Edouard, Julien Matheron, Xavier Ragot, and Juan F Rubio-Ramirez. 2017. "Precautionary saving and aggregate demand." *Quantitative Economics* 8 (2):435–478.

Chari, Varadarajan V, Patrick J Kehoe, and Ellen R McGrattan. 2007. "Business cycle accounting." *Econometrica* 75 (3):781–836.

Chien, YiLi, Harold Cole, and Hanno Lustig. 2011. "A multiplier approach to understanding the macro implications of household finance." *The Review of Economic Studies* 78 (1):199–234.

———. 2012. "Is the volatility of the market price of risk due to intermittent portfolio rebalancing?" *American Economic Review* 102 (6):2859–96.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. 2011. "When is the government spending multiplier large?" *Journal of Political Economy* 119 (1):78–121.

Constantinides, George M. 1982. "Intertemporal asset pricing with heterogeneous consumers and without demand aggregation." *Journal of business* :253–267.

Debortoli, Davide and Jordi Galí. 2017. "Monetary policy with heterogeneous agents: Insights from TANK models." Tech. rep., Unpublished manuscript.

———. 2022. "Idiosyncratic income risk and aggregate fluctuations." Tech. rep., National Bureau of Economic Research.

Den Haan, Wouter J, Pontus Rendahl, and Markus Riegler. 2017. "Unemployment (fears) and deflationary spirals." *Journal of the European Economic Association* 16 (5):1281–1349.

Fernández-Villaverde, Jesús, Juan F Rubio-Ramírez, Thomas J Sargent, and Mark W Watson. 2007. "ABCs (and Ds) of understanding VARs." *American Economic Review* 97 (3):1021–1026.

Galí, Jordi, J David López-Salido, and Javier Vallés. 2007. "Understanding the effects of government spending on consumption." *Journal of the european economic association* 5 (1):227–270.

Gertler, Mark and Peter Karadi. 2011. "A model of unconventional monetary policy." *Journal of monetary Economics* 58 (1):17–34.

Gertler, Mark and Nobuhiro Kiyotaki. 2010. "Financial intermediation and credit policy in business cycle analysis." In *Handbook of monetary economics*, vol. 3. Elsevier, 547–599.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima. 2016. "Doves for the rich, hawks for the poor? Distributional consequences of monetary policy." .

Guerrieri, Veronica and Guido Lorenzoni. 2017. "Credit crises, precautionary savings, and the liquidity trap." *The Quarterly Journal of Economics* 132 (3):1427–1467.

Gust, Christopher, Edward Herbst, David López-Salido, and Matthew E Smith. 2017. "The empirical implications of the interest-rate lower bound." *American Economic Review* 107 (7):1971–2006.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. "The Nature of Countercyclical Income Risk." *Journal of Political Economy* 122 (3):621–660.

Hagedorn, Marcus, Iourii Manovskii, and Kurt Mitman. 2019. "The fiscal multiplier." Tech. rep., National Bureau of Economic Research.

Heathcote, Jonathan and Fabrizio Perri. 2018. "Wealth and volatility." *The Review of Economic Studies* 85 (4):2173–2213.

Heathcote, Jonathan, Fabrizio Perri, and Gianluca Violante. 2010. "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States 1967-2006." *The Review of Economic Dynamics* 13 (1):15–51.

Jermann, Urban and Vincenzo Quadrini. 2012. "Macroeconomic effects of financial shocks." *American Economic Review* 102 (1):238–71.

Jones, Callum, Virgiliu Midrigan, and Thomas Philippon. 2018. "Household Leverage and the Recession." Tech. rep., National Bureau of Economic Research.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. "Monetary policy according to HANK." *American Economic Review* 32 (3):697–743.

Kaplan, Greg and Giovanni L Violante. 2014. "A model of the consumption response to fiscal stimulus payments." *Econometrica* 82 (4):1199–1239.

———. 2018. "Microeconomic Heterogeneity and Macroeconomic Shocks." *Journal of Economic Perspectives* 108 (3):167–194.

———. 2022. "The Marginal Propensity to Consume in Heterogeneous Agent Models." Tech. rep., National Bureau of Economic Research.

Kehoe, Timothy J and David K Levine. 1993. "Debt-constrained asset markets." *The Review of Economic Studies* 60 (4):865–888.

Kitagawa, Genshiro. 1996. "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models." *Journal of Computational and Graphical Statistics* 5:1–25.

Kocherlakota, Narayana and Luigi Pistaferri. 2009. "Asset pricing implications of Pareto optimality with private information." *Journal of Political Economy* 117 (3):555–590.

Koop, Gary, M Hashem Pesaran, and Simon M Potter. 1996. "Impulse response analysis in nonlinear multivariate models." *Journal of econometrics* 74 (1):119–147.

Krueger, Dirk and Hanno Lustig. 2010. "When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)?" *Journal of Economic Theory* 145 (1):1–41.

Krueger, Dirk, Hanno Lustig, and Fabrizio Perri. 2008. "Evaluating asset pricing models with limited commitment using household consumption data." *Journal of the European Economic Association* 6 (2-3):715–726.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri. 2016. "Macroeconomics and household heterogeneity." In *Handbook of Macroeconomics*, vol. 2. Elsevier, 843–921.

Krusell, Per and Anthony A Smith, Jr. 1998. "Income and wealth heterogeneity in the macroeconomy." *Journal of political Economy* 106 (5):867–896.

Lee, Seungcheol, Ralph Luetticke, and Morten O Ravn. 2021. "Financial frictions: micro vs macro volatility." .

Makowski, Louis. 1983a. "Competition and unanimity revisited." *The American Economic Review* 73 (3):329–339.

———. 1983b. "Competitive stock markets." *The Review of Economic Studies* :305–330.

Maliar, Lilia and Serguei Maliar. 2001. "Heterogeneity in capital and skills in a neoclassical stochastic growth model." *Journal of Economic Dynamics and Control* 25 (9):1367–1397.

———. 2003. "The representative consumer in the neoclassical growth model with idiosyncratic shocks." *Review of Economic Dynamics* 6 (2):362–380.

Maliar, Lilia, Serguei Maliar, and Fernando Valli. 2010. "Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm." *Journal of Economic Dynamics and Control* 34 (1):42–49.

McKay, Alisdair. 2017. "Time-varying idiosyncratic risk and aggregate consumption dynamics." *Journal of Monetary Economics* 88:1–14.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. "The power of forward guidance revisited." *American Economic Review* 106 (10):3133–58.

McKay, Alisdair and Ricardo Reis. 2016. "The role of automatic stabilizers in the US business cycle." *Econometrica* 84 (1):141–194.

Mendoza, Enrique G. 2010. "Sudden stops, financial crises, and leverage." *American Economic Review* 100 (5):1941–66.

Nakajima, Tomoyuki. 2005. "A business cycle model with variable capacity utilization and demand disturbances." *European Economic Review* 49 (5):1331–1360.

Ravn, Morten O and Vincent Sterk. 2017. "Job uncertainty and deep recessions." *Journal of Monetary Economics* 90:125–141.

———. 2021. "Macroeconomic fluctuations with HANK & SAM: An analytical approach." *Journal of the European Economic Association* 19 (2):1162–1202.

Schaab, Andreas. 2020. "Micro and macro uncertainty." Tech. rep., Unpublished manuscript.

Storesletten, Kjetil, Chris Telmer, and Amir Yaron. 2004. "Cyclical Dynamics in Idiosyncratic Labor Market Risk." *Journal of Political Economy* 112 (3):695–717.

Villalvazo, Sergio. 2021. "Inequality and Asset Prices during Sudden Stops." *Manuscript, University of Pennsylvania* .

Vissing-Jørgensen, Annette. 2002. "Limited asset market participation and the elasticity of intertemporal substitution." *Journal of political Economy* 110 (4):825–853.

Werning, Iván. 2015. "Incomplete markets and aggregate demand." Tech. rep., National Bureau of Economic Research.

Online Appendix to "Imperfect Risk Sharing and the Business Cycle"

by David Berger, Luigi Bocola and Alessandro Dovis

A RA Representation for other class of economies

In this appendix we derive the RA representation for economies not included in the class considered in the main text. These examples illustrate how our main theoretical results can be extended with few modifications to other environments. We start by adding capital accumulation to our baseline framework. We then derive the RA representation for an economy with financial frictions as in Gertler and Karadi (2011) and Gertler and Kiyotaki (2010), one for a small open economy where asset prices affect the amount of borrowing households can do, and one that features durable consumption goods. Finally, we consider a "two-agent" New Keynesian economy where a subset of agents cannot trade the risk-free bond like in Bilbiie (2008), Bilbiie (2021), and Debortoli and Galí (2017).

In particular, we show that these different models result in a different set of equations for the RA representation and in some cases introduce additional wedges, but they do not change the way we measure the preference wedges other than for the case with durable goods.

A.1 Model with capital

We now introduce capital accumulation in our baseline specification. The law of motion for aggregate capital is

$$K(z^t) \leq (1-\delta)K(z^{t-1}) + I(z^t) \tag{A.1}$$

where δ is the depreciation rate and $I(z^t)$ is the new capital produced. Capital is produced by capital good producers that make new capital using the final output and subject to adjustment costs. Capital producers discount the future dividends using the real state price $Q(z^t)$. Letting $q_0(z^t) = \prod_{z^j \preceq z^t} Q(z^j)$ be the time-0 price and $q_k(z^t)$ be the nominal price of capital, the problem for a representative capital producer is

$$\max_{I(z^t)} \sum_t \sum_{z^t} q_0(z^t) \left[\frac{q_k(z^t)}{P(z^t)} - \left(1 + f\left(\frac{I(z^t)}{I(z^{t-1})}\right)\right) \right] I(z^t) \tag{A.2}$$

where the adjustment costs are captured by the convex function f.

The capital stock is held by a representative mutual fund that maximizes:

$$\max_{K(z^t)} \sum_t \sum_{z^t} q_0(z^t)[(q_k(z^t)(1-\delta)+R(z^t))K(z^{t-1})-q_k(z^t)K(z^t)] \tag{A.3}$$

given the initial capital stock.

The monopolistic competitive firms produce the intermediate inputs using the technology $y_i\left(z^t\right)=A\left(z_t\right)k_i\left(z^{t-1}\right)^{\alpha}n_i\left(z^t\right)^{1-\alpha}$. Its problem can be split in two subproblems. First, the firm chooses the optimal input mix to minimize its marginal cost:

$$mc\left(z^t\right)=\min_{k,n}\frac{W\left(z^t\right)}{P\left(z^t\right)}n+\frac{R\left(z^t\right)}{P\left(z^t\right)}k$$

subject to

$$A\left(z_t\right)k^{\alpha}n^{1-\alpha}\geq 1.$$

Second, given the optimal factor allocation, the firm chooses its price to solve:

$$V\left(P_j,z^t\right) = \max_{p_j,y_j}\frac{p_jy_j}{P\left(z^t\right)}-mc(z^t)y_j-\frac{\kappa}{2}\left[\frac{p_j}{P_j(1+\bar{\pi})}-1\right]^2+\sum_{z^{t+1}}Q(z^{t+1}|z^t)V\left(p_j,z^{t+1}\right) \tag{A.4}$$

subject to the demand function (9).

The problem for the household and the monetary policy rule are unchanged.[28]

We have the analog of Proposition 1:

Proposition 3. *Suppose* $\left\{C(z^t),I(z^t),K(z^t),Y(z^t),\pi(z^t),q_k(z^t),i(s^t)\right\}$ *are part of an equilibrium of an heterogeneous agent economy in this section. Then, they must satisfy the aggregate Euler equation (21), the Euler equation for capital*

$$\frac{q_k(z^t)}{P(z^t)}=\sum_{z^{t+1}}Q(z^{t+1})\left[q_k(z^{t+1})(1-\delta)+\alpha\frac{Y(z^{t+1})}{K(z^t)}\right], \tag{A.5}$$

the price of capital condition

$$\begin{aligned}\frac{q_k(z^t)}{P(z^t)} &= 1+f\left(\frac{I(z^t)}{I(z^{t-1})}\right)+f'\left(\frac{I(z^t)}{I(z^{t-1})}\right)\left(\frac{I(z^t)}{I(z^{t-1})}\right)- \\ &\quad \sum_{z^{t+1}}Q(z^{t+1})f'\left(\frac{I(z^{t+1}s)}{I(z^t)}\right)\left(\frac{I(z^{t+1})}{I(z^t)}\right)^2,\end{aligned} \tag{A.6}$$

[28] Alternatively, we could have households hold capital directly.

the Phillips curve

$$\tilde{\pi}\left(z^t\right) = \frac{Y(z^t)}{\kappa\left(\mu-1\right)}\left[\mu mc(z^t)-1\right] + \sum_{z^{t+1}} Q(z^{t+1}|z^t)\tilde{\pi}\left(z^{t+1}\right), \tag{A.7}$$

the Taylor rule (14), the resource constraint

$$Y(z^t) = C(z^t) + I(z^t)\left(1+f\left(\frac{I(z^t)}{I(z^{t-1})}\right)\right) + \frac{\kappa}{2}\left[\frac{\pi(z^t)-\bar{\pi}}{1+\bar{\pi}}\right]^2, \tag{A.8}$$

and the law of motion for the capital stock (A.1), where the real state price is given by (24) and the marginal cost by

$$mc\left(z^t\right) = \left(\frac{Y\left(z^t\right)}{K\left(z^{t-1}\right)}\right)^{\alpha} \frac{\left[\omega\left(z^t\right)^{-\psi}\chi\left(\frac{Y\left(z^t\right)}{A(z_t)K\left(z^{t-1}\right)^{\alpha}}\right)^{\psi/(1-\alpha)} C\left(z^t\right)^{\sigma}\right]^{1-\alpha}}{\left(1-\alpha\right)^{1-\alpha}} \frac{1}{A\left(z_t\right)}, \tag{A.9}$$

given $\left\{\beta\left(v^t, z^{t+1}\right), \omega\left(z^t\right)\right\}$ *defined in (19) and (20) and the initial capital stock.*

The critical assumption for the equivalent representation for the model with capital is that the firm that does the capital accumulation process uses the aggregate state price (24) to discount dividends in (A.2) and (A.3). This assumption mirrors the one for the monopolistic competitive firms in (12). Our implicit assumption is that firms value dividends in a given state z^t using the valuation of the agent with the highest valuation.

A.2 Gertler-Karadi-Kiyoyaki model of financial sector

We next consider the RA-representation for a class of heterogeneous agent New Keynesian economies where the financial sector is modeled along the lines of Gertler and Karadi (2011) and Gertler and Kiyotaki (2010). In these economies, asset prices and the net worth of the financial sector matter for aggregate dynamics because of financial frictions. See Lee, Luetticke, and Ravn (2021) for a recent analysis of this type of economy.

Baseline model. Households can invest in the risk-free nominal bond, $b\left(s^t\right)$, and deposit in a bank, $d\left(s^t\right)$. The household's problem is:

$$\max_{c,l,b,d} \sum_t \sum_{s^t} \beta^t \Pr\left(s^t|s_0\right) U\left(c\left(s^t\right), l\left(s^t\right)\right)$$

subject to

$$P\left(z^t\right)c\left(s^t\right)+\frac{d\left(s^t\right)}{1+i\left(z^t\right)}+\frac{b\left(s^t\right)}{1+r\left(z^t\right)}\leq W\left(z^t\right)e\left(v_t\right)l\left(s^t\right)+d\left(s^{t-1}\right)+b\left(s^{t-1}\right)+T\left(s^t\right)$$

$$d\left(s^t\right),b\left(s^t\right)\geq 0$$

where $T\left(s^t\right)$ is a transfer function that distributes the profits of final good firms and the net dividends paid by the financial sector. Potentially, we can allow households to trade in other assets as in the text.

Financial firms (for short banks) raise deposits from the households and can invest in physical capital.[29] The amount of deposits raised from consumers is constrained by an enforcement constraint of the form

$$V\left(z^t,n\right)\geq\lambda(z^t)q\left(z^t\right)k' \tag{A.10}$$

where $V\left(z^t,n\right)$ is the value of a bank with net-worth n at the aggregate history z^t, $q(z^t)k'$ is the value of the assets and $\lambda(z^t)$ is a potentially stochastic parameter. Moreover, banks cannot issue new equity and they exit the market with some exogenous probability $1-\eta$ in any period. Exiting banks are replaced by new banks with an aggregate initial net-worth of n_0. Banks maximize the present discounted value of dividends, x, using the aggregate state price $Q\left(z^t\right)$ to discount the future. The bank problem is

$$V\left(z^t,n\right)=\max_{x,k',b',d'}x+\sum_{z_{t+1}}Q\left(z^{t+1},z^t\right)\left[\eta V\left(z^{t+1},n'\left(z^{t+1}\right)\right)+\left(1-\eta\right)n'\left(z^{t+1}\right)\right]$$

subject to

$$x+q\left(z^t\right)k'\leq n+\frac{d'}{1+r\left(z^t\right)}$$

$$n'\left(z^{t+1}\right)=\left[q\left(z^{t+1}\right)+R_k\left(z^{t+1}\right)\right]k'-d'$$

the enforcement constraint (A.10), and the inability to raise new equity, $x\geq 0$.

The representative investment and final consumption good firms solve the problems in (A.2) and (A.4). (Note that we could add a shock to the quality of capital as common in the literature about frictions in financial intermediation. We choose not for parsimony.)

As it is well known, because of the linearity of the returns, the banks' problem is linear

[29] It is without loss of generality to assume that banks do not invest in nominal bonds. In fact, if the enforcement constraint is slack then nominal bonds are perfect substitute with (negative) deposits. If the enforcement constraint is binding, banks never hold nominal bonds.

in net-worth and the problem of the financial sector admits aggregation. In particular,

$$V\left(z^t, n\right) = J\left(z^t\right) n,$$

where the value per-unit of net-worth, $J\left(z^t\right)$, solves the following functional equation:

$$J\left(z^t\right) = \max_{k', d'} \sum_{z_{t+1}} Q\left(z^{t+1}, z^t\right) \left[\eta J\left(z^{t+1}\right) n'\left(z^{t+1}\right) + (1-\eta)\, n'\left(z^{t+1}\right)\right] \tag{A.11}$$

subject to

$$\begin{aligned} q\left(z^t\right) k' &\leq 1 + \frac{d'}{1 + r\left(z^t\right)}, \\ n'\left(z^{t+1}\right) &= \left[q\left(z^{t+1}\right) + R_k\left(z^{t+1}\right)\right] k' - d', \\ J\left(z^t\right) &\geq \lambda(z^t) q\left(z^t\right) k'. \end{aligned}$$

In setting up the problem in (A.11) we used that it is (at least weakly) optimal not to pay dividends before the banks must exit. Thus, the aggregate net-worth for operating banks (surviving and new-entrants), evolves according to

$$N\left(z^t\right) = \eta \left[\left(q\left(z^t\right) + R_k\left(z^t\right)\right) K\left(z^{t-1}\right) - D\left(z^{t-1}\right)\right] + n_0 \tag{A.12}$$

and, since only exiting banks pay dividends, the net-dividends paid by banks in aggregate is

$$X\left(z^t\right) = (1-\eta) \left[\left(q\left(z^t\right) + R_k\left(z^t\right)\right) K\left(z^{t-1}\right) - D\left(z^{t-1}\right)\right] - n_0. \tag{A.13}$$

Also in this economy we have an analog to Proposition 1:

Proposition 4. *Let* $\left\{C(z^t), I(z^t), K(z^t), Y(z^t), \pi(z^t), q(z^t), i(z^t), J\left(z^t\right), D\left(z^t\right), N\left(z^t\right), r\left(z^t\right)\right\}$ *be part of an equilibrium of an heterogenous agent economy described in this section. Then, they must satisfy the aggregate Euler equation (21), the price of capital condition (A.6), the Phillips curve (A.7) the Taylor rule (14), the resource constraint (A.8), the law of motion for the capital stock (A.1),* $r\left(z^t\right) = i\left(z^t\right)$, $J\left(z^t\right)$ *solves (A.11),* $K\left(z^t\right) = k\left(z^t\right) N\left(z^t\right)$ *and* $D\left(z^t\right) = d\left(z^t\right) N\left(z^t\right)$ *where* $k\left(z^t\right)$ *and* $d\left(z^t\right)$ *are the decision rules in (A.11),* $N\left(s^t\right)$ *is given by (A.12), and where the real state prices and the marginal cost are given by (24) and (A.9), given* $\left\{\beta\left(v^t, z^{t+1}\right), \omega\left(z^t\right)\right\}$ *defined in (19) and (20), the initial capital stock and aggregate banks' net-worth.*

The proof of this proposition mirrors the one for our baseline case. The only thing to

note is that, from the household's problem, deposits and bonds are perfect substitutes so

$$\frac{1}{1+i\left(z^t\right)}=\frac{1}{1+r\left(z^t\right)}=\max_{v^t}\sum_{z^{t+1}}\beta\Pr\left(z^{t+1}|z^t,v^t\right)\hat{\beta}\left(z^{t+1},v^t\right)\frac{U_c\left(z^{t+1}\right)/P\left(z^{t+1}\right)}{U_c\left(z^t\right)/P\left(z^t\right)}.$$

Bank loans to households. Here we allow the banks to make loans to consumers as in Lee, Luetticke, and Ravn (2021).[30] In this case, we need an additional wedge to capture how consumers' borrowing crowds out investment in capital.

Let $\ell(s^t)$ denote household's bank loans and $1+r_\ell(z^t)$ the interest on such loans. The household budget constraint is

$$\begin{aligned}&P\left(z^t\right)c\left(s^t\right)+\frac{d\left(s^t\right)}{1+i\left(z^t\right)}+\frac{b\left(s^t\right)}{1+r\left(z^t\right)}-\frac{\ell\left(s^t\right)}{1+r_\ell\left(z^t\right)}\\&\leq W\left(z^t\right)e\left(v_t\right)l\left(s^t\right)+d\left(s^{t-1}\right)+b\left(s^{t-1}\right)-\ell(s^{t-1})+T\left(s^t\right)\end{aligned}$$

and the inequality constraints are

$$d\left(s^t\right),b\left(s^t\right),\ell\left(s^t\right)\geq 0,\quad \ell(s^t)\leq\bar{\ell}$$

where $\bar{\ell}$ is a debt limit. The previous example is a special case with $\bar{\ell}=0$.

The normalized bank problem can be written as

$$J\left(z^t\right)=\max_{k',d',\ell'}\sum_{z_{t+1}}Q\left(z^{t+1},z^t\right)\left[\eta J\left(z^{t+1}\right)n'\left(z^{t+1}\right)+(1-\eta)\,n'\left(z^{t+1}\right)\right] \tag{A.14}$$

subject to

$$\begin{aligned}q\left(z^t\right)k'+\frac{\ell'}{1+r_\ell(z^t)}&\leq 1+\frac{d'}{1+r\left(z^t\right)},\\n'\left(z^{t+1}\right)&=\left[q\left(z^{t+1}\right)+R_k\left(z^{t+1}\right)\right]k'+\ell'-d',\\J\left(z^t\right)&\geq\lambda(z^t)\left(q\left(z^t\right)k'+\frac{\ell'}{1+r_\ell(z^t)}\right).\end{aligned}$$

In this case, we can derive the RA representation by defining the additional wedge

$$\Psi(z^t)\equiv\sum_{v^t}\Pr(v^t|z^t)\ell(z^t,v^t) \tag{A.15}$$

[30] We could also introduce working capital loans that monopolistic competitive firms need to finance their wage bill. This would introduce an aggregate labor wedge in the model that can help to account for aggregate data as shown in Jermann and Quadrini (2012).

which is the aggregate amount of loans that banks provide in equilibrium. Moreover, the law of motion for aggregate net worth is

$$N(z^t) = \eta \left[(q(z^t) + R_k(z^t))K(z^{t-1}) + \Psi(z^{t-1}) - D(z^{t-1}) \right] + n_0. \tag{A.16}$$

We have the following result:

Proposition 5. *Let* $\{C(z^t), I(z^t), K(z^t), Y(z^t), \pi(z^t), q(z^t), i(z^t), J\left(z^t\right), D\left(z^t\right), L(z^t), N\left(z^t\right), r\left(z^t\right), r_\ell(z^t)\}$ *be part of an equilibrium of an heterogenous agent economy described in this section. Then, they must satisfy the aggregate Euler equation (21), the price of capital condition (A.6), the Phillips curve (A.7) the Taylor rule (14), the resource constraint (A.8), the law of motion for the capital stock (A.1),* $r\left(z^t\right) = i\left(z^t\right)$, $J\left(z^t\right)$ *solves (A.14),* $K\left(z^t\right) = k\left(z^t\right) N\left(z^t\right)$, $D\left(z^t\right) = d\left(z^t\right) N\left(z^t\right)$ *and* $L\left(z^t\right) = \ell\left(z^t\right) N\left(z^t\right)$ *where* $k\left(z^t\right)$, $d\left(z^t\right)$ *and* $\ell\left(z^t\right)$ *are the decision rules in (A.14),* $N\left(s^t\right)$ *is given by (A.16),* $L(z^t) = \Psi(z^t)$*, the real state prices and the marginal cost are given by (24) and (A.9), given* $\left\{\beta\left(v^t, z^{t+1}\right), \omega\left(z^t\right), \Psi(z^t)\right\}$ *defined in (19), (20) and (A.15), the initial capital stock and aggregate banks' net-worth.*

This economy shows that when assets other than the nominal risk-free bonds enters in the problem of other agents and are not perfect substitute with the nominal bonds then generally one has to introduce another wedge to obtain the RA representation.

To understand how this additional wedge affects aggregates, note that imposing market clearing $L(z^t) = \Psi(z^t)$, in the aggregate enforcement constraint we have

$$J\left(z^t\right) N(z^t) \geq \lambda(z^t) \left(q\left(z^t\right) K(z^{t+1}) + \frac{\Psi(z^t)}{1 + r_\ell(z^t)} \right)$$

Thus, an increase in the demand for bank loans tightens the banks' enforcement constraint and thus capital holding must be reduced. This has important implications. For example, an increase in idiosyncratic income risk that increases the dispersion in household's borrowing decision, is not neutral if bank's enforcement constraint is binding even if the increase in aggregate loans is compensated by an increase in aggregate deposits. This is because banks need to use scarce net-worth to back loans and must reduce capital holdings.

Moreover, this economy potentially can generate larger fluctuations in the discount factor wedge $\beta(v^t, z^{t+1})$. This is because when the banks' collateral constraint is binding, the higher costs of borrowing induces the households with negative idiosyncratic shocks to borrow less. Thus, as a tightening of the borrowing constraint studied within the class of economies in the main text, this in turn makes consumption shares more volatile and results in a higher discount factor wedge for the marginal investors.

A.3 SOE with debt limits

Next, we consider another example of an economy with collateral constraints where asset prices matter for amount of debt that can be issued by the households. Consider a SOE as in Mendoza (2010) with heterogenous agents. See Villalvazo (2021) for a recent example. Households can borrow and save in an uncontigent international bond with exogenous return $R\left(z^t\right)$ and can save in the domestic capital (land) in fixed supply. The maximal amount of debt that households can take is subject to a debt limit that depends on the value of the household's capital (land) holdings.

The household's problem is:

$$\max_{\{c,l,b,a\}} \sum_t \sum_{s^t} \beta^t \Pr\left(s^t|s_0\right) U\left(c\left(s^t\right), l\left(s^t\right)\right)$$

subject to the budget constraint

$$c\left(s^t\right) + \frac{b\left(s^t\right)}{R\left(z^t\right)} + q\left(z^t\right) a\left(s^t\right) \leq w\left(z^t\right) e\left(v_t\right) l\left(s^t\right) + b\left(s^{t-1}\right) + a\left(s^{t-1}\right)\left(q\left(z^t\right) + d\left(z^t\right)\right),$$

the debt limit

$$-\frac{b\left(s^t\right)}{R\left(z^t\right)} \leq \phi a\left(s^t\right) q\left(z^t\right),$$

and a non-negativity constraint for capital holdings, $a\left(s^t\right) \geq 0$.

Domestic firms are competitive and operate a constant return to scale technology $F(K, L)$ so

$$w\left(z^t\right) = F_l\left(L\left(z^t\right), K\right)$$
$$d\left(z^t\right) = F_k\left(L\left(z^t\right), K\right)$$

Since capital is in fixed supply, market clearing requires that

$$\sum_{v^t} \Pr\left(v^t\right) a\left(z^t, v^t\right) = K \quad \text{for all } z^t$$

while the international interest rate $R\left(z^t\right)$ is given.

This economy admits the following RA representation:

Proposition 6. *Suppose* $\left\{C\left(z^t\right), L_e\left(z^t\right), B\left(z^t\right)\right\}$ *and* $\left\{q\left(z^t\right), d\left(z^t\right), w\left(z^t\right)\right\}$ *are part of an equilibrium of an heterogenous agent economy described in this section, and for any* z^t *there is always*

an individual with a slack debt limit. Then, they must satisfy

$$\frac{1}{R\left(z^t\right)} = \max_{v^t} \sum_{z^{t+1}} \Pr\left(z^{t+1}|s^t\right) \beta\beta\left(z^{t+1}, v^t\right) \frac{U_c\left(z^{t+1}\right)}{U_c\left(z^t\right)}$$

$$-\omega\left(z^t\right) U_l\left(z^t\right) = U_c\left(z^t\right) w\left(z^t\right)$$

$$q\left(z^t\right) = \max_{v^t} \sum_{s^{t+1}} \Pr\left(s^{t+1}|s^t\right) \beta\beta\left(z^{t+1}, v^t\right) \frac{U_c\left(z^{t+1}\right)}{U_c\left(z^t\right)} \left[q\left(z^{t+1}\right) + d\left(z^{t+1}\right)\right]$$

$$w\left(z^t\right) = F_l\left(L_e\left(z^t\right), K\right)$$

$$d\left(z^t\right) = F_k\left(L_e\left(z^t\right), K\right)$$

and the consolidated budget constraint

$$C\left(z^t\right) + \frac{B\left(z^t\right)}{R\left(z^t\right)} = F\left(K, L_e\left(z^t\right)\right) + B\left(z^{t-1}\right),$$

given $\left\{\beta\left(v^t, z^{t+1}\right), \omega\left(z^t\right)\right\}$ *defined in (19) and (20), and the initial aggregate debt.*

To see this result, consider an equilibrium for the original SOE. Letting $\lambda\left(s^t\right)$ and $\mu\left(s^t\right)$ be the multipliers on the budget constraint and the debt limit, the household's focs for $b(s^t)$ and $a(s^t)$ are

$$\frac{\lambda\left(s^t\right)}{R\left(z^t\right)} = \sum_{s^{t+1}} \lambda\left(s^{t+1}\right) + \frac{\mu\left(s^t\right)}{R\left(z^t\right)}$$

$$\lambda\left(s^t\right) q\left(z^t\right) = \sum_{s^{t+1}} \lambda\left(s^{t+1}\right)\left(q\left(z^{t+1}\right) + d\left(z^{t+1}\right)\right) + \mu\left(s^t\right) \phi q\left(z^t\right)$$

where $\lambda(s^t) = \beta \Pr(s^t|s_0) u_c(s^t)$. Thus,

$$\frac{1}{R\left(z^t\right)} \geq \sum_{s^{t+1}} \beta \Pr\left(s^{t+1}|s^t\right) \frac{U_c\left(s^{t+1}\right)}{U_c\left(s^t\right)}$$

and

$$q\left(z^t\right)\left(1 - \mu\left(s^t\right)\phi\right) = \sum_{s^{t+1}} \beta \Pr\left(s^{t+1}|s^t\right) \frac{U_c\left(s^{t+1}\right)}{U_c\left(s^t\right)} \left[q\left(z^{t+1}\right) + d\left(z^{t+1}\right)\right].$$

Therefore,

$$\frac{1}{R\left(z^t\right)} = \max_{v^t} \sum_{z^{t+1}} \Pr\left(z^{t+1}|s^t\right) \beta\beta\left(z^{t+1}, v^t\right) \frac{U_c\left(z^{t+1}\right)}{U_c\left(z^t\right)}$$

and, if there is always an individual with a slack debt limit,

$$q\left(z^t\right) = \max_{v^t} \sum_{s^{t+1}} \Pr\left(s^{t+1}|s^t\right) \beta\beta\left(z^{t+1}, v^t\right) \frac{U_c\left(z^{t+1}\right)}{U_c\left(z^t\right)} \left[q\left(z^{t+1}\right) + d\left(z^{t+1}\right)\right]$$

where the wedges are defined as in the text. The derivation of the aggregate labor supply condition is the same as in the text.

Thus, we can we map the heterogeneous agent economy to a representative agent economy with the same aggregates (i.e. we get the same consumption, output, and current account here) and asset prices. The representative agent economy is a frictionless economy with no debt limits, not an economy with an aggregate debt limit (or we can construct asset prices for the marginal agent for which the constraint is not binding and so irrelevant). All the effects of binding debt limits for a fraction of the population are summarized by the discount factor wedge $\beta\left(z^{t+1}, v^t\right)$.

The last two examples illustrate how our logic extends pretty much unchanged to environments where asset prices are relevant for production or the amount of credit in the economy. Our approach would have to be amended in models of entrepreneurship where wealth and asset prices matter production choices at the household level.

A.4 Durable and Non-Durable Consumption Goods

We now consider an economy with durable goods. We show that in general we have to modify the definition of the discount factor and labor disutility wedge and add an another wedge to obtain an RA representation, but our logic still extends to this environment. We then discuss under which conditions the discount factor and labor supply wedge reduce to the ones in the text.

Suppose that households' preferences are

$$U(x, l) = \frac{x^{1-\sigma} - 1}{1-\sigma} - \chi \frac{l^{1+\psi}}{1+\psi}$$

where x is a composite of durable and non-durable consumption goods, d and c respectively. For simplicity, we let

$$x = G(c, d) = c^{\alpha} d^{1-\alpha}.$$

Durable depreciates at a rate δ. For ease of exposition, we simplify the production side of the economy by assuming that output is produced by competitive firms that produce durable and non-durable goods using labor only. Prices are flexible so we consider a

real economy. The two types of goods are perfect substitute in production. The resource constraint is

$$C\left(z^t\right) + D\left(z^t\right) \leq A\left(z_t\right) L\left(z^t\right) + (1-\delta) D\left(z^t\right), \tag{A.17}$$

where $D\left(z^t\right)$ is the aggregate stock of durable goods.

We assume that durables are illiquid in that a households cannot sell their durable stock and they can be used as collateral in the debt limit. For simplicity, we omit assets other than the risk free bond and the durable goods but they can be added as long as they satisfy the conditions in the text. The household solves

$$\max_{x(s^t),c(s^t),d(s^t),b(s^t)} \sum_t \sum_{s^t} \beta^t \Pr\left(s^t\right) U\left(x\left(s^t\right), l\left(s^t\right)\right)$$

subject to

$$x\left(s^t\right) = G\left(c\left(s^t\right), k\left(s^t\right)\right)$$

$$c\left(s^t\right) + d\left(s^t\right) + b\left(s^t\right) \leq R\left(s^t\right) b\left(s^{t-1}\right) + e\left(v_t\right) W\left(z^t\right) l\left(s^t\right)$$

$$b\left(s^t\right) \geq -\phi d\left(s^t\right)$$

$$d\left(s^t\right) \geq (1-\delta) d\left(s^{t-1}\right)$$

for some given initial stock of durables and bonds.

The solution of the household's problem is characterized by

$$U_x\left(s^t\right) G_c\left(s^t\right) + \mu\left(s^t\right) = \sum_{s^{t+1}} \beta \Pr\left(s^{t+1}|s^t\right) U_x\left(s^{t+1}\right) G_c\left(s^{t+1}\right) R\left(z^t\right) \tag{A.18}$$

$$\frac{G_c\left(s^t\right)}{G_d\left(s^t\right)} = \frac{\lambda\left(s^t\right)}{\lambda\left(s^t\right) + \mu\left(s^t\right)\phi + \kappa\left(s^t\right) - \sum_{s^{t+1} \succeq s^t} \left[\lambda\left(s^{t+1}\right) + \kappa\left(s^{t+1}\right)\right](1-\delta)} \tag{A.19}$$

$$U_x\left(s^t\right) G_c\left(s^t\right) = \frac{-U_l\left(s^t\right)}{e\left(v_t\right) W\left(z^t\right)} \tag{A.20}$$

where $\lambda\left(s^t\right)$, $\mu\left(s^t\right)$ and $\kappa\left(s^t\right)$ are the (normalized) Lagrange multipliers on the last three constraints. The intertemporal Euler equation, using our functional forms, can be written as

$$x\left(s^t\right)^{1-\sigma} / c\left(s^t\right) \geq \sum_{s^{t+1}} \beta \Pr\left(s^{t+1}|s^t\right) x\left(s^{t+1}\right)^{1-\sigma} / c\left(s^{t+1}\right) R\left(z^t\right)$$

or

$$X\left(z^t\right)^{1-\sigma} / C\left(z^t\right) \geq \sum_{z^{t+1}} \beta\tilde{\beta}\left(v^t, z^{t+1}\right) \Pr\left(z^{t+1}|z^t\right) X\left(z^{t+1}\right)^{1-\sigma} / C\left(z^{t+1}\right) R\left(z^t\right)$$

where

$$\tilde{\beta}\left(v^t, z^{t+1}\right) \equiv \sum_{v^{t+1}} \Pr\left(v^{t+1}|v^t, z^{t+1}\right) \frac{X\left(z^t\right)^{1-\sigma}/C\left(z^t\right)}{X\left(z^{t+1}\right)^{1-\sigma}/C\left(z^{t+1}\right)} \frac{x\left(z^{t+1}\right)^{1-\sigma}/c\left(s^{t+1}\right)}{x\left(s^t\right)^{1-\sigma}/c\left(s^t\right)} \tag{A.21}$$

$$= \sum_{v^{t+1}} \Pr\left(v^{t+1}|v^t, z^{t+1}\right) \left[\frac{c\left(z^t\right)/C\left(z^t\right)}{c\left(s^{t+1}\right)/C\left(z^{t+1}\right)}\right]^{-\sigma} \left[\frac{d\left(s^t\right)/c\left(s^t\right)}{d\left(s^{t+1}\right)/c\left(s^{t+1}\right)} / \frac{D\left(z^t\right)/C\left(z^t\right)}{D\left(z^{t+1}\right)/C\left(z^{t+1}\right)}\right]^{(1-\alpha)(1-\sigma)}$$

The condition (A.19) can be written as

$$\frac{d\left(s^t\right)}{c\left(s^t\right)} = \frac{1-\alpha}{\alpha} \frac{\lambda\left(s^t\right)}{\lambda\left(s^t\right) + \mu\left(s^t\right)\phi + \kappa\left(s^t\right) - \sum_{s^{t+1} \succeq s^t} \left[\lambda\left(s^{t+1}\right) + \kappa\left(s^{t+1}\right)\right](1-\delta)}$$

Thus, the durable to non-durable ratio is not always equal to $(1-\alpha)/\alpha$.

Define

$$\tilde{\omega}\left(z^t\right) \equiv \left[\sum_{v^t} \Pr\left(v^t|z^t\right) \left(\frac{x\left(s^t\right)^{1-\sigma}/c\left(s^t\right)}{X\left(z^t\right)^{1-\sigma}/C\left(z^t\right)}\right)^{1/\psi} e\left(v_t\right)^{\frac{1+\psi}{\psi}}\right]^{-\psi} \tag{A.22}$$

$$= \left[\sum_{v^t} \Pr\left(v^t|z^t\right) \left(\left[\frac{d\left(s^t\right)/c\left(s^t\right)}{D\left(z^t\right)/C\left(z^t\right)}\right]^{(1-\alpha)(1-\sigma)} \left(\frac{c\left(s^t\right)}{C\left(z^t\right)}\right)^{-\sigma}\right)^{1/\psi} e\left(v_t\right)^{\frac{1+\psi}{\psi}}\right]^{-\psi}$$

and

$$\zeta\left(z^t\right) \equiv \frac{D\left(z^t\right)}{C\left(z^t\right)} / \frac{1-\alpha}{\alpha}.$$

We have the following result:

Proposition 7. *Suppose that* $\left\{C\left(z^t\right), D\left(z^t\right), L\left(z^t\right), R\left(z^t\right)\right\}$ *are part of an equilibrium of an heterogenous agent economy described in this section. Then, they must satisfy the aggregate Euler equation,*

$$X\left(z^t\right)^{1-\sigma}/C\left(z^t\right) = \max_{v^t} \sum_{z^{t+1}} \beta\tilde{\beta}\left(v^t, z^{t+1}\right) \Pr\left(z^{t+1}|z^t\right) X\left(z^{t+1}\right)^{1-\sigma}/C\left(z^{t+1}\right) R\left(z^t\right)$$

the labor supply condition

$$A\left(z_t\right) X\left(z^t\right)^{1-\sigma}/C\left(z^t\right) = \tilde{\omega}\left(z^t\right) \chi L\left(z^t\right)^{\psi}$$

the resource constraint (A.17), and

$$\frac{D\left(z^t\right)}{C\left(z^t\right)} = \frac{1-\alpha}{\alpha}\zeta\left(z^t\right),$$

where $X\left(z^t\right) = C\left(z^t\right)^{\alpha} D\left(z^t\right)^{1-\alpha}$, *given* $\left\{\tilde{\beta}\left(v^t, z^{t+1}\right), \tilde{\omega}\left(z^t\right), \zeta(z^t)\right\}$.

The proposition highlights that an additional wedge, $\zeta(z^t)$, is needed in the RA representation to allow the model to match the ratio between durable and non-durable consumption. This is because different binding patterns for the collateral constraint and the non-negative of durable purchases can induce different ratios in the cross-section. For the very same reason, the discount factor and the labor disutility wedges have slightly different definitions as shown in (A.21) and (A.22). For instance, the discount factor wedge $\tilde{\beta}(v^t, z^{t+1})$ is the expectations over the product of two terms: the relative growth rate of non-durable consumption shares – as in the case in the text – and the growth rate of the durable to non-durable consumption relative to the aggregate. This second term modifies the definition of the discount factor wedge.

The wedges in (A.21) and (A.22) reduce to the ones in the baseline case considered in the text when $c(s^t)/d(s^t)$ is constant in the cross-section. In this case, we have

$$\tilde{\beta}\left(v^t, z^{t+1}\right) = \sum_{v^{t+1}} \Pr\left(v^{t+1}|v^t, z^{t+1}\right)\left[\frac{c\left(z^t\right)/C\left(z^t\right)}{c\left(s^{t+1}\right)/C\left(z^{t+1}\right)}\right]^{-\sigma} = \hat{\beta}\left(v^t, z^{t+1}\right)$$

$$\tilde{\omega}\left(z^t\right) = \left[\sum_{v^t} \Pr\left(v^t|z^t\right)\left(\frac{c\left(s^t\right)}{C\left(z^t\right)}\right)^{-\sigma/\psi} e\left(v_t\right)^{\frac{1+\psi}{\psi}}\right]^{-\psi} = \omega\left(z^t\right)$$

and of-course $\zeta\left(z^t\right)$ is constant. Thus, our economy with non-durable consumption only approximates well the economy with durables if the cross-sectional variance of $c(s^t)/d(s^t)$ is sufficiently small.

A.5 TANK

The class of economies considered in the main text does not include economies where a set of agents must pay a fixed cost to trade the nominal risk-free bond. Thus, the class of models considered does not include limited participation models. (Note that all the other assets other than the risk-free nominal bond can be subjected to fixed costs and limited participation. For instance, agents may have to pay a fixed cost in order to trade shares of firms.) Here we show that a similar RA-representation as the one in Proposition 1 is valid

for these model. The only difference is that one cannot identify the marginal agents by choosing the one with the highest discount factor wedge from micro data. This is because the presence of a fixed cost implies that the marginal agent is not necessarily the one with the highest marginal incentive to save. Thus, to make our method operational, a researcher must know which agents are marginal.

Consider an economy where the macro block is the same as in the main text. Assume that there are two types of households: a *participant*, indexed by super-script p, and a *non-participant*, np, with measure μ_p and μ_{np} respectively. Assume there are no idiosyncratic shocks. The budget constraint of a participant is

$$\begin{aligned} &P\left(z^t\right) c_p\left(z^t\right) + \frac{b_p\left(z^t\right)}{1+i\left(z^t\right)} + a_p\left(z^t\right) q\left(z^t\right) \\ &\leq W\left(z^t\right) e_p\left(z_t\right) l_p\left(z^t\right) + T_p\left(z^t\right) + b_p\left(z^t\right) + a_p\left(z^{t-1}\right)\left(R_k\left(z^t\right) + q\left(z^t\right)\right) \end{aligned}$$

where $a_p\left(z^t\right)$ are the participant holdings of the shares of monopolistic competitive firms, $p\left(z^t\right)$ is the price of such share, and $R\left(z^t\right)$ are the dividends distributed by the firms. The budget constraint of a non-participant is

$$P\left(z^t\right) c_{np}\left(z^t\right) \leq W\left(z^t\right) e_{np}\left(z_t\right) l_{np}\left(z^t\right) + T_{np}\left(z^t\right).$$

Thus, non-participants are hand-to-mouth agents. Market clearing requires that $a_p\left(z^t\right) = 1$ and $b_p\left(z^t\right) = B\left(z^t\right)$.

A possible foundation for this economy is that there are heterogenous fixed costs in accessing assets markets: participants have a zero cost while non-participants face a cost sufficiently high to induce them not to participate.

Toward establishing the RA representation, define the following statistics

$$\begin{aligned} \beta_p(z^{t+1}) &\equiv \left(\frac{c_p(z^{t+1})/C(z^{t+1})}{c_p(z^t)/C(z^t)}\right)^{-\sigma} = \left(\frac{1/\left[\mu_p + \mu_{np} c_{np}\left(z^{t+1}\right)/c_p\left(z^{t+1}\right)\right]}{1/\left[\mu_p + \mu_{np} c_{np}\left(z^t\right)/c_p\left(z^t\right)\right]}\right)^{-\sigma} \\ \omega\left(z^t\right) &\equiv \left[\sum_{i=p,np} \mu_i \left(\frac{c_i(z^t)}{C(z^t)}\right)^{\frac{-\sigma}{\psi}} e_i\left(z_t\right)^{\frac{1+\psi}{\psi}}\right]^{-\psi}. \end{aligned}$$

where ω is just a specialized version of ω in the text. We then have the following proposition:

Proposition 8. *Suppose that* $\left\{C\left(z^t\right), Y\left(z^t\right), \pi\left(z^t\right), i(z^t)\right\}$ *are part of an equilibrium for the*

TANK economy. Then, they must satisfy the aggregate Euler equation,

$$\frac{1}{1+i\left(z^{t}\right)}=\beta\sum_{z_{t+1}}\left\{\frac{\Pr\left(z^{t+1}|z^{t}\right)\theta(z_{t+1})\beta_{p}\left(z^{t+1}\right)}{1+\pi\left(z^{t+1}\right)}\left(\frac{C\left(z^{t+1}\right)}{C\left(z^{t}\right)}\right)^{-\sigma}\right\}, \tag{A.23}$$

the Phillips curve,

$$\tilde{\pi}\left(z^{t}\right)=\frac{Y(z^{t})}{\kappa\left(\mu-1\right)}\left[\mu\frac{\chi Y\left(z^{t}\right)^{\psi}C\left(z^{t}\right)^{\sigma}\omega\left(z^{t}\right)}{A(z_{t})^{1+\psi}}-1\right]+\sum_{z^{t+1}}\beta_{p}(z^{t+1})\tilde{\pi}\left(z^{t+1}\right) \tag{A.24}$$

the Taylor rule (14), the resource constraint

$$Y(z^{t})=C(z^{t})+\frac{\kappa}{2}\left[\frac{\pi(z^{t})-\bar{\pi}}{1+\bar{\pi}}\right]^{2}, \tag{A.25}$$

given $\left\{\beta_{p}\left(z^{t+1}\right),\omega\left(z^{t}\right)\right\}$.

This is exactly the same result we have in Proposition 1. The only difference is that it may not be true that $\beta_{p}\left(z^{t+1}\right)\geq\beta_{np}\left(z^{t+1}\right)$ for all z^{t+1}. Thus, we cannot identify the participants (the marginal agents) directly from individual consumption data only.

B Derivations for the economy in Section 3.4

Here we provide the detailed derivations for the example 3 in Section 3.4. In period 1, since we assume that $T_{1H}=(1+i)\frac{P_0}{P_1}B/\mu_H$, agents are homogeneous and consumption is determined by the static first order condition

$$w_1\frac{1}{C_1}=w_1\frac{1}{w_1 l_1}=\chi l_1^{\psi}.$$

Under the simplifying assumption that $\psi=1$ and $\chi=1$, consumption equals 1 in period 1 because from the firm's problem $w_1=1$. Thus, we know that in period 1:

$$c_{0H}=c_{1H}=C_1=l_{1H}=l_{1L}=1.$$

Given an interest rate i, the equilibrium objects in period 0, c_{0H}, c_{0L}, l_{0H}, l_{0L}, w_0, C_0, solve the following system of equations:

$$\frac{1}{c_{0H}}=\beta\left(1+i\right)\frac{P_0}{P_1}\frac{1}{c_{1H}}$$

$$c_{0H} = e_H w_0 l_{0H} - \frac{T - \Pi}{\mu_H} + T$$

$$c_{0L} = e_L w_0 l_{0L} + T$$

$$e_i w_0 \frac{1}{c_{0i}} = \chi l_{0i}^{\psi}$$

$$C_0 = \sum_i \mu_i e_i l_{0i}$$

$$C_0 = \sum_i \mu_i c_{0i}$$

Under the assumption that monetary policy sets the nominal rate so that the real rate is $1/\beta$, $\beta(1+i)\frac{P_0}{P_1} = 1$, the system simplifies to

$$\frac{1}{c_{0H}} = 1 \Rightarrow c_{0H} = 1$$

$$c_{0H} = e_H w_0 l_{0H} - \frac{T - fp}{\mu_H} + T$$

$$c_{0L} = e_L w_0 l_{0L} + T$$

$$l_{0H} = e_H w_0$$

$$l_{0L} = e_L w_0 \frac{1}{e_L w_0 l_{0L} + T} \Rightarrow -e_L w_0 l_{0L}^2 - T l_{0L} + e_L w_0 = 0 \Rightarrow l_{0L} = \frac{\sqrt{T^2 + 4(e_L w_0)^2} - T}{2 e_L w_0}$$

and

$$\sum_i \mu_i e_i l_{0i} = \sum_i \mu_i c_{0i} \tag{A.26}$$

where

$$Y_0 = C_0 = \sum_i \mu_i c_{0i}.$$

We can express (A.26) as

$$\mu_H e_H^2 w_0^2 + \mu_L \frac{\sqrt{T^2 + 4(e_L w_0)^2} - T}{2} = \mu_H w_0 + \mu_L w_0 \left(\frac{\sqrt{T^2 + 4(e_L w_0)^2} - T}{2} + T \right) \tag{A.27}$$

which implicitly defines the real wage as a function of T and μ_L, $w_0(T, \mu_L)$. Given $w_0(T, \mu_L)$, the other equilibrium allocations are given by

$$\begin{aligned} l_{0H} &= e_H w_0(T, \mu_L) \\ l_{0L} &= \frac{\sqrt{T^2 + 4\left(e_L w_0(T, \mu_L)\right)^2} - T}{2 e_L w_0(T, \mu_L)} \\ c_{0H} &= 1 \\ c_{0L} &= e_L w_0(T, \mu_L) l_{0L} + T \end{aligned}$$

and

$$C_0 = Y_0 = \sum_i \mu_i c_{0i}.$$

The equations above fully characterize the equilibrium outcome. Thus, to solve for the equilibrium outcome given T and μ_L we simply need to numerically solve (A.27). To generate Figure 1 in the text, we compute the numerical derivative of the equilibrium output with respect to T for different levels of μ_L.

C Monte Carlo analysis

In this appendix we study how well our approximations for estimating the realization of the wedges from panel data and approximating their stochastic processes work in practice by performing a Monte Carlo analysis on data simulated from the Guerrieri and Lorenzoni (2017) and the Krusell and Smith (1998) economy.

C.1 The Guerrieri and Lorenzoni (2017) economy

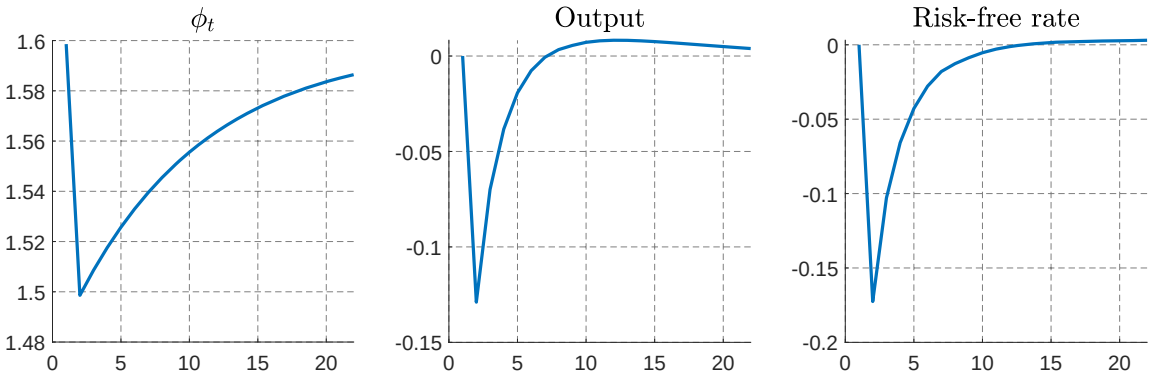
We focus on the flexible price economy ($\kappa = 0$) where households have the following preferences over consumption and labor,[31]

$$U(c, l) = \frac{c^{1-\sigma} - 1}{1 - \sigma} + \chi \frac{(1 - l)^{1-\psi}}{1 - \psi}.$$

Households produce a consumption good using the technology,

$$y_{i,t} = e_{i,t} l_{i,t},$$

[31] The functional form for labor disutility here is different from the one we considered in Section 2. Because of that, the expressions for ω_t and the labor supply condition in the RA representation are slightly different from those of Proposition 1.



Notes: *The figure reports the response of output and the risk-free rate after a tightening of the borrowing limit. Output is reported in percentage deviations from its steady state while the risk-free rate in percentage points.*

Figure A-1: IRFs to a tightening of the borrowing limit

where $e_{i,t}$ is an idiosyncratic shock to labor productivity. The households can only save and borrow in a real non-contingent bond earning the risk-free real rate R_t and they face a borrowing limit $b_{i,t+1} \geq -\phi_t$. The debt limit ϕ_t is the only aggregate shock in this economy, and it follows the AR(1) process,

$$\phi_t = (1-\rho)\bar{\phi} + \rho\phi_{t-1} + \sigma_\phi \varepsilon_t.$$

In equilibrium, the households' net demand of bonds equal the supply of bonds by the government, B. We follow Guerrieri and Lorenzoni (2017) in the calibration of households' preferences, the Markov process $e_{i,t}$, the bond supply B and the average borrowing limit $\bar{\phi}$. We set $\rho = 0.90$ and $\sigma_\phi = 0.05$. Since we consider a flexible price economy, we can abstract from the specification of monetary policy and nominal variables. We perform simulations by applying the methodology of Boppart, Krusell, and Mitman (2018).

Figure A-1 plots the IRFs to a reduction in ϕ_t of two standard deviations—a tightening of the borrowing limit. We can see that the shock leads to a persistent decline in output and in the risk-free rate. The decline in aggregate output is due to a change in the composition of labor supply. When the borrowing limit tightens, borrowers cut current consumption while savers increase it. Due to wealth effects, the former increase their labor supply while the latter reduce it. Because borrowers have on average a lower realization of $e_{i,t}$ than savers, these changes in the composition of the labor force lead to a decline in average productivity and thus in output. The decline in the risk-free rate is due to the decline in the supply of bonds issued by borrowers.

The RA representation. If aggregate output, Y_t, consumption, C_t, and the risk-free real rate, R_t, are part of an equilibrium of the Guerrieri and Lorenzoni (2017) economy, then,

they satisfy the following conditions

$$\begin{aligned} \frac{1}{R_t} &= \beta \max_i \mathbb{E}_t \left\{ \beta_{i,t+1} \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \right\}, \\ C_t^{-\sigma} &= \omega_t \left[1 - Y_t \right]^{-\psi} \\ Y_t &= C_t \end{aligned} \tag{A.28}$$

where $\beta_{i,t+1}$ is defined in equation (19) and ω_t is given by

$$\omega_t = \left\{ \mathbb{E}_t \left[\left(\frac{\varphi_{i,t}^{\frac{\sigma}{\psi}} e_{i,t}^{\frac{\psi-1}{\psi}}}{\chi^{-\frac{1}{\psi}}} \right) \mathbb{1}(l_{i,t} > 0) \right] - \frac{\mathbb{E}_t[e_{i,t} \mathbb{1}(l_{i,t} > 0)] - 1}{C_t^{\frac{\sigma}{\psi}}} \right\}^{\psi}. \tag{A.29}$$

This is the analog of Proposition 1 for this economy. The first equation in (A.28) is the aggregate Euler equation for bonds. The second equation is obtained from aggregating the labor supply conditions of households that are working, and using the production function to substitute hours worked in efficiency units for aggregate output. This is the analog of the Phillips curve when there are no nominal rigidities, $\kappa = 0$. The third equation is the resource constraint.

The variable ω_t in equation (A.29) represents the time-varying disutility of labor in the RA representation and it captures the compositional changes in labor supply that take place in the heterogenous agent economy. This expression is a generalization of equation (20) that allows for movements of labor at the extensive margin, which occur in the Guerrieri and Lorenzoni (2017) due to their preference specifications.[32] Because there are no price rigidities, aggregate output and consumption are determined using just the last two equations in (A.28), and they are only functions of ω_t. Holding ω_t constant, fluctuations in $\mathbb{E}_t[\beta_{i,t+1}]$ do not have any effect on output, and are fully reflected on the risk-free rate.

In a recursive competitive equilibrium of the heterogeneous agent economy, the exogenous aggregate state is $z_t = z_t^{HA} = [\phi_t]$, and there are no exogenous states that are directly part of the RA representation, $z_t^{RA} = \emptyset$. The endogenous aggregate state is $X_t = X_t^{HA} = \Psi_t(e, b)$ where Ψ_t is the joint distribution of individual productivities and of asset holdings, and there are no endogenous states that are directly part of the RA representation, $X_t^{RA} = \emptyset$. So, in the Guerrieri and Lorenzoni (2017) economy no state-variable of the heterogeneous agent economy appears in the RA representation, and equation (32)

[32] It is straightforward to verify that the two expressions coincide when $\mathbb{1}(l_{i,t} > 0) = 1$ for all i, t.

Table A-1: Monte Carlo analysis

	Panel (a): Stochastic process for the wedges						
	$\Phi_{\beta,\beta}$	$\Phi_{\beta,\omega}$	$\Phi_{\omega,\beta}$	$\Phi_{\omega,\omega}$	$\sigma_{\varepsilon_\beta} \times 100$	$\sigma_{\varepsilon_\omega} \times 100$	$\rho_{\varepsilon_\beta,\varepsilon_\omega}$
MC Average	0.00	0.60	-0.02	0.54	0.43	0.31	-0.92
5^{th} Percentile	-0.15	0.40	-0.13	0.36	0.38	0.27	-0.94
95^{th} Percentile	0.15	0.81	0.10	0.68	0.50	0.36	-0.88
	Panel (b): Business cycle statistics						
	Mean(Y_t)	Mean(r_t)	Stdev(y_t)	Stdev(r_t)	Acorr(y_t)	Acorr(y_t)	Corr(y_t, r_t)
HA economy	0.42	0.63	0.08	0.11	0.53	0.58	0.98
MC Average	0.42	0.62	0.08	0.10	0.56	0.50	0.98
5^{th} Percentile	0.42	0.61	0.06	0.07	0.40	0.10	0.89
95^{th} Percentile	0.42	0.63	0.09	0.13	0.68	0.75	0.99

Notes: Panel (a) reports the estimates of a VAR(1) for $\mathbf{T}_{t+1}$. The first row reports the average of the coefficients across the Monte Carlo replications, the remaining two columns the 5^{th} and 95^{th} percentile. Panel (b) reports moments for output and the risk-free rate in the heterogeneous agent economy and in the RA representation. Y_t is output in level while y_t is output in percentage deviation from the ergodic mean. The risk-free rate is reported in percentages (annualized).

specializes to the VAR process

$$\mathbf{T}_{t+1} = \mathbf{\Phi}(\mathbf{L}) \times \mathbf{T}_t + \varepsilon_{t+1}, \tag{A.30}$$

Monte Carlo analysis. We simulate 500 panel datasets from the original heterogeneous agent economy, with each panel having 10000 households and 100 quarters. We compute $\beta_{g,t+1}$ for the four households' groups using equation (28), and we compute ω_t using equation (A.29). In our simulations, the high-income/low net worth and high income/high net worth groups have on average the highest $\beta_{g,t+1}$ in our simulations, suggesting that those households are typically the ones on their Euler equation. We include in $\mathbf{T}_{t+1}$ only $\beta_{g,t+1}$ for the high income/high net worth group in order the minimize the number of parameters to be estimated in the law of motion for $\mathbf{T}_{t+1}$.

For each panel dataset, we estimate the VAR process in equation (A.30), setting the lag structure to 1. Panel (a) of Table A-1 reports the estimates of the VAR process. We can see that ω_t is on average positively autocorrelated ($\phi_{\omega,\omega} = 0.54$), and it predicts high values for $\beta_{i,t+1}$, as $\phi_{\beta,\omega} = 0.61$. Given the estimated parameters, we solve for the policy functions of the RA representation and simulate aggregate variables. Panel (b) of Table A-1 reports first and second moments for output and the risk-free rate in those simulations. Specifically, it reports for each moment the average across these Monte Carlo replications along with the 5^{th} and 95^{th} percentile. The raw "HA economy" reports Monte Carlo average of the

same moments computed from simulations of the actual heterogeneous agents economy. We can see that the RA representation reproduces very accurately the underlying stochastic behavior of output and the risk-free rate in the heterogeneous agent economy.

This analysis shows that our approach to compute the discount factor wedge and the proposed approximation to the law of motion for the wedges work extremely well in this application.

C.2 The Krusell and Smith (1998) economy

The heterogenous agents economy. We consider an economy with capital accumulation and flexible prices, $\kappa = 0$. Households inelastically supply labor, $\chi = 0$, and their idiosyncratic labor productivity follows a Markov process and can take on two values, $e \in \{1, 0\}$. The households decide how much to consume and save, and financial markets are incomplete in that households can only trade claims on the capital stock subject to a debt limit.

The representative firm produces a final good using the technology

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha},$$

where A_t is an aggregate TFP shock, K_t is the capital stock, and L_t is total worked hours worked in efficiency units. Given the process for idiosyncratic risk, the latter equals to

$$L_t = 1 - u_t,$$

where u_t is the fraction of agents that are currently unemployed.

We let A_t to take two values, $A_L < 1 < A_H$. In addition, we allow for the aggregate shock to affect the distribution of idiosyncratic labor productivities: a low A_t is associated with a higher probability that a household samples $e = 0$, which makes idiosyncratic risk countercyclical.[33]

We consider two calibrations. The first is the one in Krusell and Smith (1998). It is well know that with those parameters' values the model delivers "approximate aggregation", in the sense that the cyclical behavior of output, consumption and investment in the incomplete market economy closely mirror those in the corresponding economy with complete financial markets—the representative-agent real business cycle model. The second, which we label "high risk calibration", is identical to the one in Krusell and Smith (1998) with

[33] Specifically, let $\pi_{AA'ee'}$ be the probability of $A_{t+1} = A'$ and $e_{i,t+1} = e'$ given that $A_t = A$ and $e_{i,t} = e$. These transition probabilities are such that u_t is only a function of A_t and can thus take two values, $u(A_L) = u_H$ and $u(A_H) = u_L$, with $u_H > u_L$.

the exception that the probability of being unemployed when $A_t = A_L$ is 20% instead of 10%.[34] This makes idiosyncratic income risk more countercyclical and, as we will see below, it breaks approximate aggregation. We numerically solve the heterogeneous agent economy using the software provided by Maliar, Maliar, and Valli (2010).

The circled lines in Figure A-2 reports the impulse response functions (IRFs) to a negative TFP shock.[35] Panel (a) reports the IRFs with the Krusell and Smith (1998) calibration: consumption, investment and output fall after the shock, and we know from their analysis that these magnitudes are essentially identical to those of the corresponding representative agent economy. Panel (b) reports the IRFs in the high risk calibration. Relative to the calibration of Krusell and Smith (1998), consumption falls by more and investment falls by less after the shock. This difference is due to the higher incidence of precautionary savings: in the high risk calibration, households have more incentives to save after a negative TFP shock because of the higher probability of being unemployed, and these precautionary motives depress aggregate consumption and increase aggregate investment, as households in this economy can save only by accumulating claims on the capital stock.

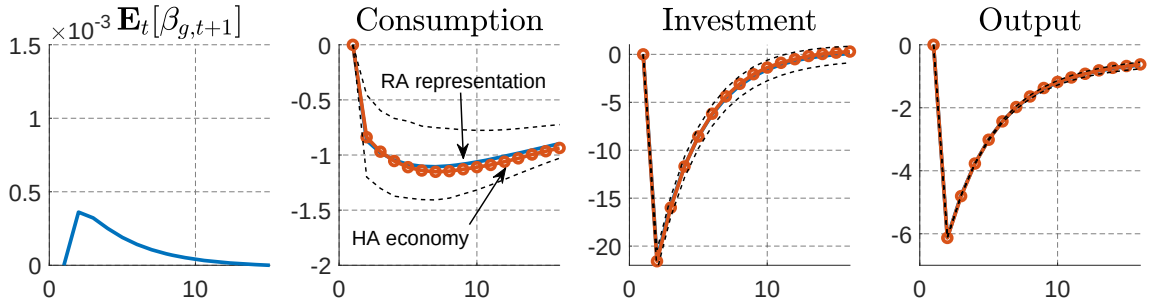
The RA representation. We now describe the RA representation of this economy. In what follows, it is convenient to define $\tilde{A}_t = A_t(1-u_t)^{1-\alpha}$. The exogenous aggregate state for this economy is $z_t = z_t^{RA} = \tilde{A}_t$ and there are no exogenous shocks that are not part of the RA representation, $z_t^{HA} = \emptyset$. The endogenous aggregate state is $X_t = [X_t^{RA}, X_t^{HA}]$ where X_t^{RA} is the level of the capital stock K_t and X_t^{HA} is the joint distribution of asset holdings and individual productivities Ψ_t. If one follows the common practice of recording only the mean capital stock as a statistic for the distribution of asset holdings then there is no problem of a missing state variable because $K_t = X_t^{RA}$.

Given that labor is inelastically supplied, there is no labor supply condition of households, so $\mathbf{T}_{t+1} = [\hat{\beta}_{1,t+1}, \ldots, \hat{\beta}_{i,t+1}]'$ does not contain ω_t. Equation (32) then specializes to

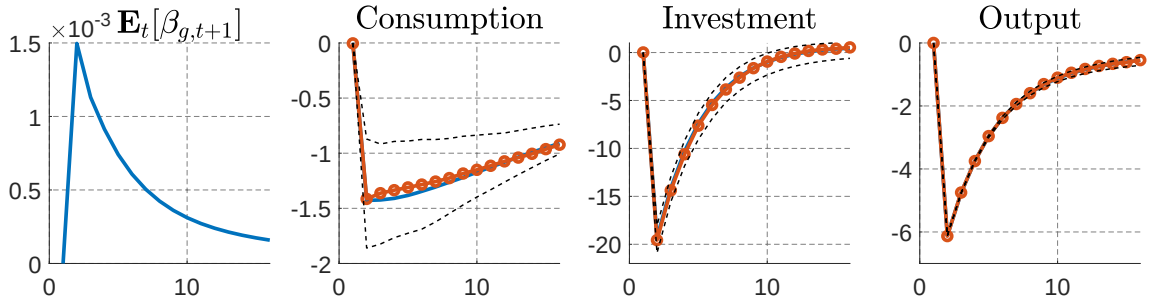
$$\mathbf{T}_{t+1} = \mathbf{\Phi}(\mathbf{L}) \times \mathbf{T}_t + \mathbf{A} \times \hat{K}_t + \mathbf{B} \times \hat{\tilde{A}}_t + \mathbf{C} \times \hat{\tilde{A}}_{t+1} + \varepsilon_{t+1}. \tag{A.31}$$

[34] When doing so, we normalize A_L so that $A_t(1-u_t)^{1-\alpha}$ is the same across the two calibrations.

[35] We compute non-linear IRFs following Koop, Pesaran, and Potter (1996). Starting from the ergodic mean of the model, we compute $2 \times M$ simulations for aggregate consumption, investment and output of length T. In the first M simulations, we restrict TFP at $t = 1$ to equal A_H. In the second set of simulations, we restrict TFP at $t = 1$ to equal A_L. To obtain the IRFs, we average the first and second sets of simulations across M and take the difference between the two paths.



(a) Krusell and Smith calibration



(b) High Risk calibration

Notes: *The circled line reports IRFs to a negative technology shock in the Krusell and Smith (1998) economy. The solid line reports the Monte Carlo average of the IRFs to a negative technology shock in the RA representation, while the dotted line reports the 5^{th} and 95^{th} percentile across the Monte Carlo simulations. Panel (a) reports this experiment under the standard calibration of the Krusell and Smith (1998), while panel (b) reports the same information for the high risk calibration. Consumption, investment and output are reported in percentage changes from their ergodic men value.*

Figure A-2: IRFs to a negative TFP shock

So, the RA representation is

$$\begin{aligned} Y_t &= K_{t+1} + C_t \\ Y_t &= \tilde{A}_t K_t^{\alpha} \\ 1 &= \beta \max_i \mathbb{E}_t \left\{ \beta_{i,t+1} \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left[\alpha \frac{Y_{t+1}}{K_{t+1}} + (1-\delta) \right] \right\}, \end{aligned} \tag{A.32}$$

subject to the stochastic process for $\tilde{A}_t$ and $\mathbf{T}_t$. This is a representative agent real business cycle model with a time-varying discount factor for the stand-in household, and can be easily solved numerically.

Monte Carlo analysis. We now study how well our procedure approximates the IRFs reported in Figure A-2. For that purpose, we proceed as follows. We simulate 500 panel datasets containing households' level information on consumption, income and assets from the original heterogeneous agent economy. Each panel has 10000 households, and lasts 100

Table A-2: Summary statistics for $\beta_{g,t+1}$

	Krussel and Smith calibration			High Risk calibration		
	Mean$(\tilde{\beta}_{g,t+1})$	Mean$(E_{g,t+1})$	R^2	Mean$(\tilde{\beta}_{g,t+1})$	Mean$(E_{g,t+1})$	R^2
$y_{L,t}/n_{L,t}$	0.004	-0.060	0.032	-0.032	0.082	0.030
$y_{L,t}/n_{H,t}$	0.015	-0.015	0.031	0.029	0.035	0.030
$y_{H,t}/n_{L,t}$	0.015	-0.014	0.031	0.079	-0.008	0.029
$y_{H,t}/n_{H,t}$	0.015	-0.014	0.032	0.078	-0.006	0.028

Notes: For each calibration, the first column reports the Monte Carlo average of the sample mean of $\tilde{\beta}_{g,t+1} \equiv (\beta_{g,t+1} - 1) \times 100$, the second column the Monte Carlo average of the sample mean of $EE_{g,t+1}$ defined in equation (A.33), and the third column reports the Monte Carlo average of the R^2 of an OLS regression of $\tilde{A}_t$ and K_t on $EE_{g,t+1}$.

quarters. For each of these panel datasets, and for each t, we partition households in four groups based on their labor income income, $y_{i,t}$, and net worth, $n_{i,t}$, as described in Section 4.1 and compute $\beta_{g,t+1}$ using equation (28). We next estimate the stochastic process (A.31), solve for the policy functions of the RA representation, and compute the IRFs to a negative TFP shock.

Let us start by studying how the $\beta_{g,t+1}$ varies across households. The first column of Table A-2 reports the Monte Carlo average of the sample mean of $\beta_{g,t+1}$ for each group. We can see that high income households have, on average, a higher $\beta_{g,t+1}$ than households with lower labor income, especially in the high-risk calibration. This is the result of two forces in the model. First, consumption shares are positively related to idiosyncratic productivity shocks because idiosyncratic risk is not perfectly insured. Second, these productivity shocks are mean-reverting. Thus, the consumption shares of households that are hit by a positive idiosyncratic shock today fall on average between today and tomorrow, explaining why this group has a higher $\beta_{g,t+1}$ on average.

To verify whether the aggregate Euler equation in (A.32) holds when using the measured wedges, we compute the following statistic

$$EE_{g,t+1} = 100 \times \left(\frac{1}{\beta} - \beta_{g,t+1} \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} R_{k,t+1} \right), \tag{A.33}$$

where $R_{k,t+1}$ is the realized return to capital in period $t+1$. By construction $EE_{g,t+1}$ should be on average equal to zero and orthogonal to the current information set for the households that are on their Euler equation—those achieving the "max" in the Euler equation of the RA representation. In the second column of Table A-2 we report the Monte Carlo average of the sample mean of $EE_{g,t+1}$, while the third column reports the Monte Carlo

average of the R^2 of an OLS regression of $\tilde{A}_t$ and K_t on $EE_{g,t+1}$. We can see that $EE_{g,t+1}$ associated to high income households are close to zero on average, and they are not predicted by the aggregate state variables of the model. These results show that our approach for measuring the wedge, despite the approximations discussed in Section 4.1, works well in this economy.

After retrieving the time path for the wedges, and for each Monte Carlo replication, we estimate the stochastic process in equation (A.31) with one lag in the autoregressive component. We include in $\mathbf{T}_t$ only the $\beta_{g,t}$ for the high income/high net worth group, as the results in Table A-2 shows that the aggregate Euler equation holds when using this wedge.[36] Given the estimated parameters, we next solve for the policy function of the equivalent RA representation in (A.32) and compute IRFs to a negative TFP shock. As we repeat this process for every Monte Carlo replication, we obtain a distribution of IRFs.

Figure A-2 reports the results: the solid line represents the average across the Monte Carlo replications while the dotted lines report the 5^{th} and 95^{th} percentile. We can see that in both calibrations the RA representation does remarkably well in retrieving the true underlying IRFs of the heterogeneous agent economy. In the Krussel and Smith calibration, $\mathbb{E}_t[\beta_{g,t+1}]$ does not change much in response to the TFP shock. That is, the IRFs in the heterogeneous agent economy are very close to those of a representative agent economy with a time-invariant discount factor. This is an implication of approximate aggregation. In the high risk calibration, instead, $\mathbb{E}_t[\beta_{g,t+1}]$ increases substantially when TFP falls. That is, the RA representation captures the countercyclical precautionary motives that arise in the original heterogeneous agent economy through an increase in the discount factor of the stand-in household.

D Data

In this appendix, we provide full details about how our empirical sample was constructed. We also benchmark our baseline sample to government aggregates and show that our sample is consistent with other aggregate trends.

D.1 Aggregate data

Real gross domestic product [GDPC1]. *U.S. Bureau of Economic Analysis*, Billions of chained 2012 dollars, Seasonally adjusted annual rate. We de-trend this series by estimating the

[36] That is to say, households that have above median income and above median net worth are most of the time unconstrained in this economy, so they achieve the "max" in the Euler equation (A.32). Results are virtually identical when we include in $\mathbf{T}_t$ the $\beta_{g,t+1}$ for the four groups.

following regression with ordinary least squares

$$\log(Y_t) = a_1 + a_2 t + a_3 t^2 + \varepsilon_{y,t},$$

where $\log(Y_t)$ is the logarithm of real gross domestic product and t is calendar time. The residual of this regression, $\varepsilon_{y,t}$ is the de-trended output series that we use in our application.

Effective Federal Funds Rate [FEDUFUNDS]. *Board of Governors of the Federal Reserve System*, Percent, Annual averages of daily figures.

Inflation [FEDUFUNDS]. *World Bank*, Percent, Annual percentage change in the Consumer Price Index.

PCE Consumption [Sum of Nondurables and Services from Table 2.3.5.]. *NIPA*, Billions of Nominal Dollars. Annual averages of quarterly figures.

Income [Sum of wages, business income, rental and asset income from NIPA Table 2.1]. *NIPA*, Billions of Nominal Dollars, Annual averages of quarterly figures.

Hours [AVHWPEUSA065NRUG]. *Penn World Tables and FRED*. Hours, Not Seasonally Adjusted.

D.2 Definition of variables and sample selection in the CE

Consumption expenditures. Our measure of consumption expenditure is close to the NIPA definition of nondurable and services expenditures. It is constructed by aggregating up the following expenditure sub-categories: food, tobacco, domestic services, adult and child care, utilities, transportation, pet expenses, apparel, education, work-related and training, healthcare, insurance, furniture rental and small textiles, housing related expenditures excluding rent.

Total hours worked. We compute total hours worked for the head of household by multiplying the number of weeks worked full or part time over the last year (*INCWEEK1*) multiplied by the numbers of hours usually worked per week (*INC_HRS1*). We obtain the same indicator for the spouse (*INCWEEK2* and *INC_HRS2*) and add the two.

Labor income. We compute labor income as the sum of total household (combine unit CU) income from earnings before taxes (pre-2004 *FSALARYX*; after 2004 *FSALARYM*), plus

the total income received from farm (pre-2004 - *FFRMINCX*; 2004-2012 *FFRMINCM*) and nonfarm business (pre-2004 *FNONFRMX*; 2004-2012 *FNONFRMM*). For 2013 onward, these two variables are subsumed into *FSMPFRMX*, which measures pre-tax labor income at the household level.

Liquid assets. It includes the total amount the households held in savings accounts in financial institutions (*SAVACCTX*), checking and brokerage accounts (*CKBKACTX*). In the CE, these amounts are only reported in the last interview. Thus they represent end of period values for the household. In order to define beginning of period values for these assets, we use the following variables (*COMPSAVX* and *COMPCKGX*), which report the total change in savings and checking accounts over the previous year, respectively. Then beginning of period values are defined as end of period values minus the change in value. From 2013 onward, we use the variable (*LIQUDYRX*) to measure liquid assets.

Illiquid Assets. It includes the value of owned automobiles (*NETPURX*), residential housing (*PROPVALX*), U.S. savings bonds (*USBNDX*), the value of all securities directly held by the household (include stocks, mutual funds and non U.S. savings bonds) (*SECESTX*), and money owned to the household by individuals outside of the household (*MONYOWDX*). The value of U.S. savings bonds and total securities are only reported in household interview. In order to define beginning of period values for these assets, we use the following variables (*COMPBNDX* and *COMPSECX*), which report the total change in U.S. savings bonds and all securities over the previous year, respectively. Then beginning of period values are defined as end of period values minus the change in value. From 2013 onward, the value of bonds and securities is captured by the variable *STOCKYRX*.

Total assets. It is the value of liquid assets plus illiquid assets each household owns.

Liabilities. It is the current value of the household's home mortgage (*QBLNCM3X*) plus the outstanding principal balance on auto debt. (*QBALNM3X*).

Net worth. Net Worth is total assets minus liabilities.

The baseline sample includes all households where the head of the household is between the ages of 22 and 64 over the period 1992-2017. We only use households who participate in all four interviews in the CE. We restrict the sample to those which the CE labels as "complete income reporters," whenever possible (the variable REPSTAT is only available until 2013), corresponding to households with at least one non-zero response to any of the income and benefits questions, those with non-negative consumption, income and hours,

those with positive food expenditures and those with non-missing information on income, net worth, wages and consumption. We use the assigned "replicate" or sample weights, designed to map the CE into the national population in all calculations. We use the CPI-U to express all monetary variables in constant 2000 dollars. To eliminate outliers and mitigate any impact of time-varying top-coding, we i) eliminate extreme observations for $c_{i,t+1}/c_{i,t}$ when computing the discount factor wedge (as in Vissing-Jørgensen (2002), we drop observations where semi-annual change in consumption share is less than 0.2 or larger than 5) and ii) winsorize the top and bottom 1% of consumption shares and relative wage observations when computing the wedges. Finally, we drop households for which we are unable to perform the Mincer regression described in the main text because of missing data on the covariates.

Taking 2006 as the year of reference, we have 5131 households that report full consumption information in all four interviews. We next keep households whose head is in the age bracket 22-64, leaving us with 3920 households that reported income and consumption in 2006. Within this group, we keep households that are considered "full income responders" (3441), and drop any household that observed a change in family size between the first and the last interview (3398). We then keep only households with information on consumption, net worth, income and hours in each year, leaving us with 2907 households for 2006. Finally, we are not able to run our Mincer regressions for households with missing co-variates. This leaves us with 2327 households in 2006.

D.3 Summary statistics

Table A-3 reports selected households' characteristics for 2006. In the CE, the average age for the head of household was 44 years, and roughly 37% of the households' head held a college degree. The average size of the household was 2.75. On average, households spent roughly 10512 dollars per person in non-durables and services, and the average income per person was 28774 dollars. Households worked 1331 hours per year per person on average, earning an average wage of 28.80 dollars per hour. The mean net worth for the household was 172963 dollars, with 16574 dollars in liquid assets. As a comparison with previous papers, the average characteristics of the household in our sample are very close to those reported in Heathcote and Perri (2018). See Table 1 in their paper.

D.4 Trends in the data

In this section we examine whether the dynamics of aggregate consumption, income, and total hours per capita in our cross sectional data capture the broad contours of national

Table A-3: Average characteristics of households in 2006

	CE
Age of head	44.32
Household size	2.75
Head with college (%)	36.87
Consumption expenditures per person	10511.51
Labor income per person	28774.22
Disposable income per person	28672.91
Hours worked per person	1330.98
Wage per hour	28.80
Household's net worth	172962.70
Liquid assets	16573.65

Notes: *The sample size is 2327 households. All statistics are computed using sample weights. All monetary variables are expressed in 2000 U.S. dollars.*

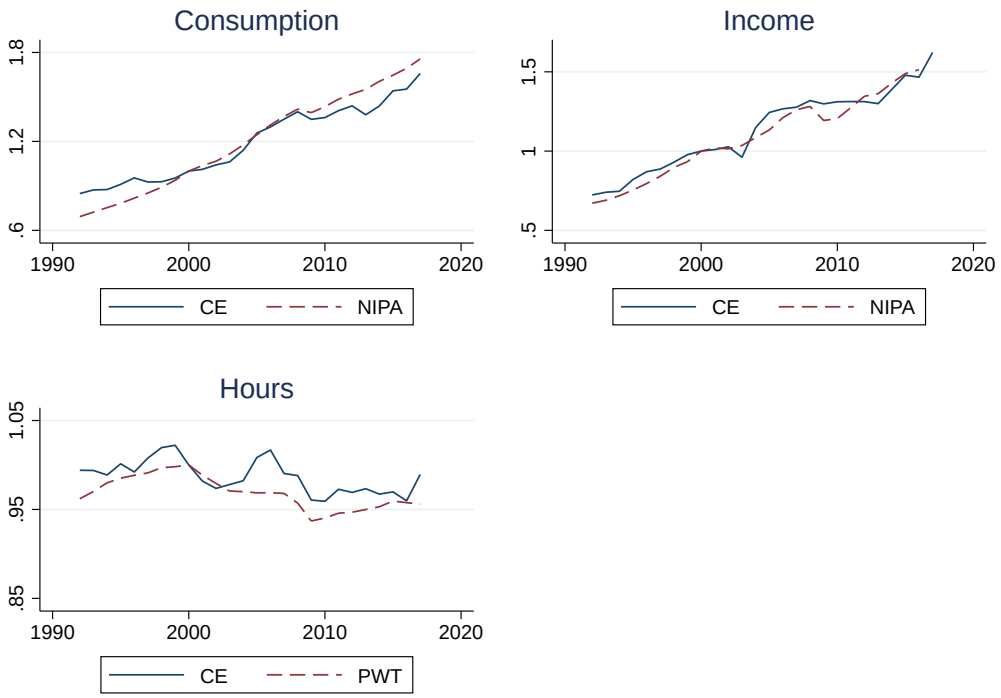
income and product accounts (NIPA) aggregates. Each series is normalized to 1 in 2000 because the levels vary somewhat across data sets. See Section D.1 for the precise variable definitions of these aggregates. These aggregates were chosen to be as close as possible to their CE equivalents. The results are shown in figure A-3.

The top left panel of figure A-3 shows the dynamics of average per capita expenditures in the CE and the closest equivalent measure in the NIPA (PCE non-durable and services expenditure). The top left panel shows average per capita income in the CE and NIPA, where our NIPA measure is the sum of wages, business income, rental and asset income from Table 2.1). The bottom panel shows average annual worked per capita in the CE as well as its aggregate counterpart obtained from the Penn World Table. While the fit is not perfect, it is clear that the dataset captures the broad contour of each aggregate series during the Great Recession.[37]

D.5 Alternative measurement of the discount factor wedge

Figure A-4 shows the results of our baseline measurement of the (high income, low net worth) $\beta_{g,t}$ where we grouped households according to whether their income at date $t-1$ is above or below median income and, within each of these two groups, weather the level of their net worth is above or below the group median. Figure A-4 reports results under the following different assumptions: when deflating consumption internally using the CE data

[37] Our figure A-3 is very similar to the relevant panels in Figure 13 of Heathcote and Perri (2018) giving us further confidence.



Notes: *Top panel compares consumption and income per capita in the CE and in NIPA. The bottom panel shows the behavior of average annual hours worked per capita in the CE and Penn World Tables (PWT). Each series is normalized to 1 in 2000.*

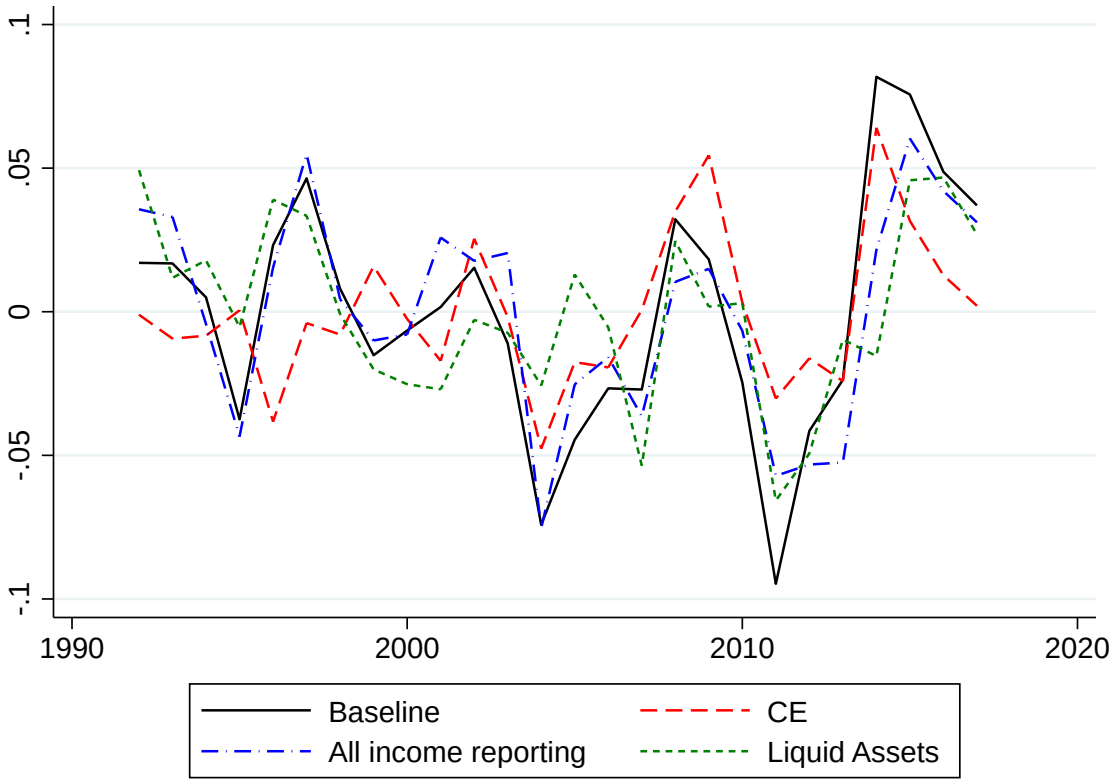
Figure A-3: Aggregate consumption, income and worked hours in the CE

rather than NIPA aggregates (red dashed line), when keeping all households who report income rather than complete income reporters (blue dotted line) and when we partition households using two income and two liquid asset groups (green dash-dotted line). Our baseline group is shown in the solid black line.

Deflating by the CE rather than NIPA mainly affects the level of $\beta_{g,t}$, but there is still a large increase during the Great Recession. Relaxing our assumptions about the reporting of income or changing the partition from net worth to liquid wealth has little effect on the main result. This suggests that our results are robust to using other natural partitions of the CE data.

E Numerical solution

Let the state vector be $\mathbf{S}_t = [i_{t-1}, \hat{\theta}_t, \hat{A}_t, \varepsilon_{m,t}, \hat{\beta}_t, \hat{\omega}_t]$. The equilibrium conditions of the model can be summarized by the following equations



Notes: *This figure shows an estimate of the $\beta_{g,t}$ for the high income/low net worth group in four different cases: our baseline with two income and two net worth groups (solid black line), when computing consumption shares using the CE data rather than NIPA for aggregate consumption (dashed red line), when keeping all households who report income rather than complete income reporters (blue dotted line), and when using net liquid assets rather than net worth when partitioning the households (green dash-dotted line).*

Figure A-4: β measurement robustness

$$Y(\mathbf{S}_t) = C(\mathbf{S}_t) + \frac{\kappa}{2}\left(\frac{\pi(\mathbf{S}_t) - \pi^\star}{1+\pi^\star}\right)^2 \tag{A.34}$$

$$Y^{\text{pot}}(\mathbf{S}_t) = \left[\frac{\exp\{\hat{A}_t\}^{1+\psi}}{\exp\{\hat{\omega}_t\}\mu\chi}\right]^{\frac{\psi}{1+\psi\sigma}} \tag{A.35}$$

$$1 + i(\mathbf{S}_t) = \max\left\{(1+i_{t-1})^{\rho_i}\left[(1+\bar{\imath})\left(\frac{1+\pi(\mathbf{S}_t)}{1+\pi^*}\right)^{\psi_\pi}\left(\frac{Y(\mathbf{S}_t)}{Y^{\text{pot}}(\mathbf{S}_t)}\right)^{\phi_y}\right]^{(1-\rho_i)}, 1\right\} \tag{A.36}$$

$$1 = [1+i(\mathbf{S}_t)]\beta\mathbb{E}_t\left[\exp\{\hat{\theta}_{t+1}+\hat{\beta}_{t+1}\}\left(\frac{C(\mathbf{S}_{t+1})}{C(\mathbf{S}_t)}\right)^{-\sigma}\frac{1}{1+\pi(\mathbf{S}_{t+1})}\right] \tag{A.37}$$

$$\begin{aligned}\frac{\pi(\mathbf{S}_t)-\pi^*}{1+\pi^*}\frac{1+\pi(\mathbf{S}_t)}{1+\pi^*} &= \frac{1}{\kappa(\mu-1)}Y(\mathbf{S}_t)\left(\mu\chi\frac{Y(\mathbf{S}_t)^\psi C(\mathbf{S}_t)^\sigma\exp\{\hat{\omega}_t\}}{\exp\{\hat{A}_t\}^{1+\psi}}-1\right)\\ &+ \beta\mathbb{E}_t\left[\exp\{\hat{\theta}_{t+1}+\hat{\beta}_{t+1}\}\left(\frac{C(\mathbf{S}_{t+1})}{C(\mathbf{S}_t)}\right)^{-\sigma}\frac{\pi(\mathbf{S}_{t+1})-\pi^*}{1+\pi^*}\frac{1+\pi(\mathbf{S}_{t+1})}{1+\pi^*}\right].\end{aligned} \tag{A.38}$$

Given policy functions for $C(\mathbf{S}_t)$ and $\pi(\mathbf{S}_t)$, we can use equations (A.34)-(A.36) to solve for $Y(\mathbf{S}_t)$ and $i(\mathbf{S}_t)$. Thus, the numerical solution of the model can be equivalently expressed as approximating $C(\mathbf{S}_t)$ and $\pi(\mathbf{S}_t)$.

Due to the max operator in equation (A.36), $C(\mathbf{S}_t)$ and $\pi(\mathbf{S}_t)$ may have kinks in a region of $\mathbf{S}_t$ where the zero lower bound constraint starts binding, a feature that makes it challenging to approximate these functions with smooth polynomials. We approach this feature following Gust, Herbst, López-Salido, and Smith (2017). Specifically, we approximate these variables using a piece-wise smooth function,

$$x(\mathbf{S}_t) = \mathbb{1}(1 + \tilde{i}(\mathbf{S}_t) > 1)\gamma_x^{\text{no zlb}\prime}\mathbf{T}(\mathbf{S}_t) + \mathbb{1}(1 + \tilde{i}(\mathbf{S}_t) \leq 1)\gamma_x^{\text{zlb}\prime}\mathbf{T}(\mathbf{S}_t), \tag{A.39}$$

where $x = \{C, \pi\}$, $1 + \tilde{i}(\mathbf{S}_t)$ is the "notional" interest rate at $\mathbf{S}_t$ (the first term inside the max operator of equation (A.36)), $\mathbf{T}(\mathbf{S}_t)$ is a vector collecting Chebyshev's polynomials evaluated at $\mathbf{S}_t$ and $\{\gamma_x^{\text{no zlb}}, \gamma_x^{\text{zlb}}\}$ a set of coefficients.

The numerical solution of the model consists in choosing $\Gamma = \{\gamma_x^{\text{no zlb}}, \gamma_x^{\text{zlb}}\}_{x=C,\pi}$ so that equations (A.37) and (A.38) are satisfied for a set of collocation points $\tilde{\mathbf{S}}^{\mathbf{i}} \in \mathcal{S}$. The choice of collocation points and the associated Chebyshev's polynomials follows the method of Smolyak. Conditional expectations in equations (A.37) and (A.38) are evaluated using Gauss-Hermite quadrature.

The algorithm for the numerical solution of the model is as follows:

Step 0.A: Defining the grid and the polynomials. Set upper and lower bounds on the state variables $\tilde{\mathbf{S}} = [i, \hat{\theta}, \hat{A}, \varepsilon_m, \hat{\beta}, \hat{\omega}]$. Given these bounds, construct a Smolyak grid and the associated Chebyshev's polynomials.

Step 1: Equilibrium conditions at the candidate solution. Start with a guess for the model's policy functions Γ^c. For each $\tilde{\mathbf{S}}^{\mathbf{i}}$, compute $C^{\text{no zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$, $C^{\text{zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$, $\pi^{\text{no zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$ and $\pi^{\text{zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$ using the coefficients in Γ^c. Evaluate equation (A.34) using $C^{\text{no zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$ and $\pi^{\text{no zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$ to obtain $Y^{\text{no zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$, and similarly obtain a value for $Y^{\text{zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$. Use equation (A.35) and (A.36) along with $Y^{\text{no zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$ and $\pi^{\text{no zlb}}(\tilde{\mathbf{S}}^{\mathbf{i}})$ to obtain the notional interest rate $1 + \tilde{i}(\tilde{\mathbf{S}}^{\mathbf{i}})$. Compute the actual interest rate $1 + i(\tilde{\mathbf{S}}^{\mathbf{i}}) = \max\{1 + \tilde{i}(\tilde{\mathbf{S}}^{\mathbf{i}}), 1\}$.

Step 3: Evaluate residual equations. For each $\tilde{\mathbf{S}}^{\mathbf{i}}$, compute the residual equations

$$\begin{aligned}
\mathcal{R}^1(\tilde{\mathbf{S}}^i)) &\equiv \left[\frac{1}{1 + \tilde{i}(\tilde{\mathbf{S}}^i)}\right] - \beta\mathbb{E}\left[\exp\{\theta' + \beta'\}\left(\frac{C(\mathbf{S}')}{C^{\text{no zlb}}(\tilde{\mathbf{S}}^i)}\right)^{-\sigma}\frac{1}{1 + \pi(\mathbf{S}')}\right] \\
\mathcal{R}^2(\tilde{\mathbf{S}}^i)) &\equiv 1 - \beta\mathbb{E}\left[\exp\{\theta' + \beta'\}\left(\frac{C(\mathbf{S}')}{C^{\text{zlb}}(\tilde{\mathbf{S}}^i)}\right)^{-\sigma}\frac{1}{1 + \pi(\mathbf{S}')}\right].
\end{aligned}$$

Similarly, compute $\mathcal{R}^3(\tilde{\mathbf{S}}^i)$ and $\mathcal{R}^4(\tilde{\mathbf{S}}^i)$ using equation (A.38).

Step 4: Iteration. Let $\mathcal{R}(\Gamma^c)$ the vector collecting all the computed residuals at the collocation point, and let r be its Euclidean norm. If $r \leq 10^{-10}$, stop the algorithm. If not, update the guess and repeat Step 1-4. □

The specifics for the algorithm are as follows. The bounds on $[\hat{\theta}, \hat{A}, \varepsilon_m, \hat{\beta}, \hat{\omega}]$ are +/- 3 standard deviations from their mean. The bounds on i is set to $[0, 0.20]$, wide enough to span the ergodic distribution of nominal interest rates. We consider a second-order Smolyak grid, and use 243 points for Gauss-Hermite quadrature (three points for each shock and tensor multiplications). Finally, we use a Newton algorithm to find the zeros of $\mathcal{R}(\Gamma^c)$ at the collocation points.

F Quantitative analysis

In this Appendix we present additional details regarding the quantitative experiments of Section 5. We start with the estimation of the model, present some indicators of model fit and finally discuss in details the counterfactual of Section 5.4.

F.1 Model estimation

Draws from the posterior distribution of the model parameters are generated using the random walk Metropolis Hastings described in An and Schorfheide (2007). The proposal distribution is a multivariate normal, with variance-covariance matrix given by $c\Sigma$, where Σ is the negative of the inverse hessian of the log-posterior evaluated at the posterior mode and c is a constant that we set to obtain roughly a 30% acceptance rate in Markov chain. We generate 2 Markov chains of 300,000 each, and discard the first 200,000 draws in each chain. The statistics of the posterior distribution of model parameters reported in Table A-4 and Table A-5 are computed by combining the last 100,000 draws for each chain.

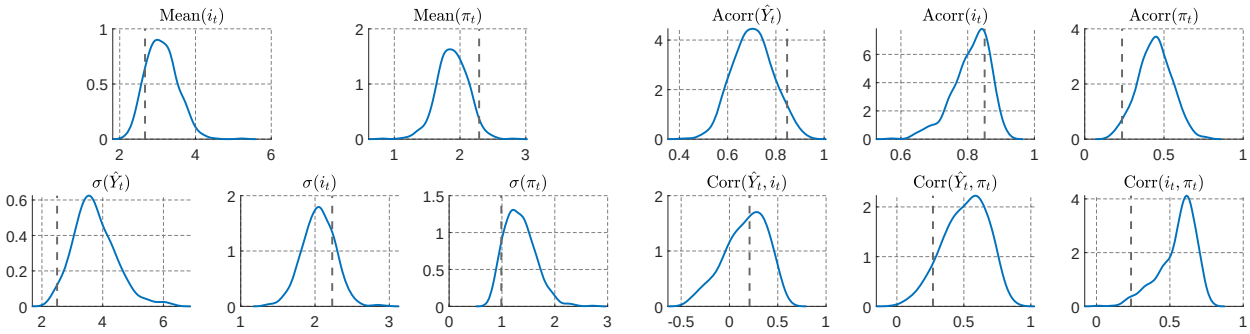
F.2 Posterior predictive checks

We perform posteriod predictive checks in order to assess model fit. Let $\{\Theta_i\}_{i=1}^N$ be $N = 1000$ draws from the posterior distribution of the model paramters. For each draw i, we solve the RA representation at Θ_i, simulate a sample of lenght $T = 1250$ of aggregate data from the RA representation, and use the last 250 draws to compute a set of statistics on the model simulated data, $\mathbf{s}^T(\Theta_i)$. In our application, we focus on the sample mean, standard deviation and autocorrelation of output, inflation and nominal interest rates, as well as

Table A-4: Prior and posterior distribution: structural parameters

		Prior		Posterior	
Parameter	Distribution	Mean	St. dev.	Mean	90% Interval
$4 \times \kappa$	Gamma	85.00	15.00	81.39	[65.92, 97.58]
ρ_i	Beta	0.50	0.25	0.64	[0.47, 0.78]
γ_π	Normal	3.00	2.00	3.88	[2.59, 5.52]
ρ_θ	Beta	0.50	0.28	0.79	[0.63, 0.93]
ρ_A	Beta	0.50	0.28	0.89	[0.78, 0.97]
$100 \times \sigma_\theta$	InvGamma	1.00	5.00	1.62	[0.90, 2.78]
$100 \times \sigma_A$	InvGamma	1.00	5.00	0.94	[0.68, 1.26]
$100 \times \sigma_m$	InvGamma	1.00	5.00	1.43	[1.07, 1.93]

Notes: *The posterior statistics reports the mean, fifth and ninety-fifth percentile of the posterior distribution estimated by pooling 2 Markov chains with 100,000 draws each (with a 200,000 draw burn-in period for each chain).*



Notes: *The solid line reports the density of* $\{\mathbf{s}^T(\Theta_i)\}_{i=1}^{1000}$ *estimated using kernel smoothing. See the text for details on the computation of* $\{\mathbf{s}^T(\Theta_i)\}_{i=1}^{1000}$*. The dotted vertical line reports the same statistic computed using U.S. data.*

Figure A-5: Posterior predictive checks

the cross-correlation among these three variables. This procedure allow us to obtain the posterior distribution of the sample statistics for the estimated model and to assess whether the model can replicate the observed statistical features for these variables. See Aruoba, Bocola, and Schorfheide (2017) for a discussion of posterior predictive analysis for the evaluation of dynamic stochastic general equilibrium models.

The solid line in Figure A-5 plots the posterior density of the sample statistics under consideration, while the dotted line reports the same statistics computed using the U.S. data in our sample. We can see that the model for the most part captures the statistical behavior of these three series: with few exceptions, each of the sample statistic under consideration lies well within the model posterior distribution for the same statistic.

Table A-5: Prior and posterior distribution: stochastic process of the wedges

		Prior		Posterior	
Parameter	Distribution	Mean	St. dev.	Mean	90% Interval
$\phi_{\beta,\beta}$	Beta	0.50	0.25	0.32	[0.15, 0.48]
$\phi_{\beta,\omega}$	Beta	0.00	0.10	-0.01	[-0.10, 0.08]
$\phi_{\omega,\beta}$	Normal	0.00	0.10	0.07	[-0.03, 0.17]
$\phi_{\omega,\omega}$	Normal	0.50	0.25	0.91	[0.81, 0.98]
$A_{\beta,i}$	Normal	0.00	0.10	0.06	[-0.07, 0.18]
$B_{\beta,\theta}$	Normal	0.00	0.10	0.01	[-0.10, 0.10]
$B_{\beta,A}$	Normal	0.00	0.10	-0.00	[-0.10, 0.10]
B_{β,ε_m}	Normal	0.00	0.10	-0.07	[-0.19, 0.05]
$C_{\beta,\theta}$	Normal	0.00	0.10	0.07	[-0.05, 0.18]
$C_{\beta,A}$	Normal	0.00	0.10	-0.04	[-0.19, 0.10]
C_{β,ε_m}	Normal	0.00	0.10	0.04	[-0.07, 0.15]
$A_{\omega,i}$	Normal	0.00	0.10	-0.01	[-0.16, 0.14]
$B_{\omega,\theta}$	Normal	0.00	0.10	0.06	[-0.11, 0.22]
$B_{\omega,A}$	Normal	0.00	0.10	0.09	[-0.04, 0.23]
B_{ω,ε_m}	Normal	0.00	0.10	0.06	[-0.07, 0.20]
$C_{\omega,\theta}$	Normal	0.00	0.10	0.04	[-0.10, 0.19]
$C_{\omega,A}$	Normal	0.00	0.10	-0.02	[-0.18, 0.16]
C_{ω,ε_m}	Normal	0.00	0.10	0.00	[-0.10, 0.11]
$100 \times \sigma_\beta$	InvGamma	1.00	5.00	3.35	[2.52, 4.33]
$100 \times \sigma_\omega$	InvGamma	1.00	5.00	2.54	[1.67, 3.62]
$\rho_{\varepsilon_\beta,\varepsilon_\omega}$	Normal	0.00	0.20	0.09	[-0.08, 0.27]
$\phi_{\omega^{\text{cm}},\omega^{\text{cm}}}$	Beta	0.50	0.25	0.81	[0.62, 0.95]
$A_{\omega^{\text{cm}},i}$	Normal	0.00	0.10	-0.05	[-0.24, 0.17]
$B_{\omega^{\text{cm}},\theta}$	Normal	0.00	0.10	-0.07	[-0.21, 0.08]
$B_{\omega^{\text{cm}},A}$	Normal	0.00	0.10	0.08	[-0.06, 0.21]
$B_{\omega^{\text{cm}},\varepsilon_m}$	Normal	0.00	0.10	0.02	[-0.12, 0.22]
$C_{\omega^{\text{cm}},\theta}$	Normal	0.00	0.10	-0.10	[-0.23, 0.01]
$C_{\omega^{\text{cm}},A}$	Normal	0.00	0.10	0.04	[-0.07, 0.16]
$C_{\omega^{\text{cm}},\varepsilon_m}$	Normal	0.00	0.10	-0.07	[-0.21, 0.06]
$100 \times \sigma_{\omega^{\text{cm}}}$	InvGamma	1.00	5.00	5.18	[3.36, 7.42]

Notes: *This table reports estimates for the parameters in (32). The parameter $A_{y,x}$ stands for the loading of an element x of X_t on the variable y in $\mathbf{T}_{t+1}$. These parameters define the matrix $\mathbf{A}$ in (32). The parameters $B_{y,x}$ and $C_{y,x}$ are defined in a similar manner. The posterior statistics are constructed as explained in the note to Table A-4.*

F.3 Counterfactuals

We now detail the counterfactual experiment of Section 5. We first explain how we use the particle filter to obtain an estimate of the structural shocks. Next, we discuss how we generate the decomposition of Figure 4.

The RA representation implies the following law of motion for the model's variables

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{g}(\mathbf{S}_t) + \eta_t \\ \mathbf{S}_t &= \mathbf{f}(\mathbf{S}_{t-1}, \varepsilon_t). \end{aligned}$$

The first set of equations describes the evolution of the observables $\mathbf{Y}_t$, with η_t being a vector of iid Gaussian errors with a diagonal variance-covariance matrix equal to $\mathbf{H}$. The second equation describes the evolution of the state variables $\mathbf{S}_t$. The vector ε_t collects the innovations to the structural shocks $\hat{\theta}_t$, $\hat{A}_t$, and $\varepsilon_{m,t}$ and the preference wedges $\hat{\beta}_t$, $\hat{\omega}_t$ and $\hat{\omega}_t^{\text{cm}}$. The functions $g(.)$ and $f(.)$ are generated using the numerical algorithm described previously and they depend implicitly on the structural parameters of the model.

Fix the vector of model parameters φ, and denote by $\mathbf{Y}^t = [\mathbf{Y}_1, \ldots, \mathbf{Y}_t]$ the vector collecting data and by $p(\mathbf{S}_t|\mathbf{Y}^t)$ the conditional distribution of the state vector given observations up to period t. Although the conditional density of $\mathbf{Y}_t$ given $\mathbf{S}_t$ is known and Gaussian, there is no analytical expression for the density $p(\mathbf{S}_t|\mathbf{Y}^t)$. We use the particle filter to approximate this density for each t. The approximation is done via a set of pairs $\{\mathbf{S}_t^i, \tilde{w}_t^i\}_{i=1}^N$, in the sense that

$$\frac{1}{N}\sum_{i=1}^{N} f(\mathbf{S}_t^i)\tilde{w}_t^i \overset{\text{a.s.}}{\to} \mathbb{E}[f(\mathbf{S}_t)|\mathbf{Y}^t].$$

We refer to $\mathbf{S}_t^i$ as a particle and to $\tilde{w}_t^i$ as its weight. The algorithm used to approximate $\{p(\mathbf{S}_t|\mathbf{Y}^t)\}_t$ builds on Kitagawa (1996), and it goes as follows:

Step 0: Initialization. Set $t = 1$. Initialize $\{\mathbf{S}_0^i, \tilde{w}_0^i\}_{i=1}^N$ and set $\tilde{w}_0^i = 1 \ \forall i$.

Step 1: Prediction. For each $i = 1, \ldots, N$, obtain a realization for the state vector $\mathbf{S}_{t|t-1}^i$ given $\mathbf{S}_{t-1}^i$ by simulating the model forward.

Step 2: Filtering. Assign to each particle $\mathbf{S}_{t|t-1}^i$ the weight

$$w_t^i = p(\mathbf{Y}_t|\mathbf{S}_{t|t-1}^i)\tilde{w}_{t-1}^i.$$

Step 3: Resampling. Rescale the weights $\{w_t^i\}$ so that they add up to one, and denote these rescaled values by $\{\tilde{w}_t^i\}$. Sample N values for the state vector with replacement from $\{\mathbf{S}_{t|t-1}^i, \tilde{w}_t^i\}_{i=1}^N$, and denote these draws by $\{\mathbf{S}_t^i\}_i$. Set $\tilde{w}_t^i = 1 \ \forall i$. If $t < T$, set

$t = t + 1$ and go to Step 1. If not, stop. □

In our exercise, the measurement equation includes nominal interest rates, linearly detrended real GDP, inflation, $\hat{\beta}_t$, $\hat{\omega}_t$ and $\hat{\omega}_t^{\text{cm}}$. The variance on the measurement errors on the first three variables is set to 0.5% of their unconditional variance, while we set the variance of the measurement errors on the wedges to 10% of their unconditional variance. We set N to 1,000,000. Given the vector of model parameters φ, we solve the model using the algorithm in Online Appendix E, apply the particle filter to our data, and obtain an estimate for the latent states. We repeat this procedure for 1000 approximately iid draws from the posterior distribution of the model parameters. Figure 3 reports the posterior mean and 90% credible sets for $\mathbf{Y}_t$ and $\mathbf{S}_t$.

In order to generate the counterfactual of Figure 4, we first solve the model setting $\hat{\beta}_t = 0$ and by setting the stochastic process of the labor supply wedge to those of ω_t^{cm}. Let's denote by $\mathbf{g}^{\text{cm}}(.)$ and $\mathbf{f}^{\text{cm}}(.)$ the policy function of the complete market counterfactual. We then compute the counterfactual value of a variable y_t as

$$y_t^{\text{cm}} = \sum_{i=1}^{N} g_y^{\text{cm}}(\mathbf{S}_t^i)\tilde{w}_t^i,$$

where $\mathbf{S}_t^i = [\hat{\theta}_t^i, \hat{A}_t^i, \varepsilon_{m,t}^i, \hat{\omega}_t^i \times (\hat{\omega}_t^{\text{cm}}/\hat{\omega}_t)]$. We repeat this procedure for 1000 draws from the posterior distribution of model parameters in order to construct credible sets.

F.4 Robustness

In our analysis we have set the population mean of the two wedges to 1. This assumption has important implications for how imperfect risk sharing affects the *average* behavior of macroeconomic variables, as it implies that the steady state would not change if we were to switch to an economy with complete markets. While our paper does not focus on long run averages, it is nonetheless important to ask whether this assumption also impacts the business cycle properties of the counterfactual economy with complete markets and, if so, whether that affects the main conclusion of the paper.

To address this question, we repeat our counterfactual analysis under different assumptions regarding the mean of the wedges. Let $\mu_{\hat{\beta}}$ be the population mean of the discount factor wedge.[38] The analysis in the main text is conducted under the assumption that $\mu_{\hat{\beta}} = 1.00$. In this Appendix, we repeat the main counterfactual experiment of the paper

[38] We will focus on the discount factor wedge because the behavior of the labor disutility wedge is not that important in our application.

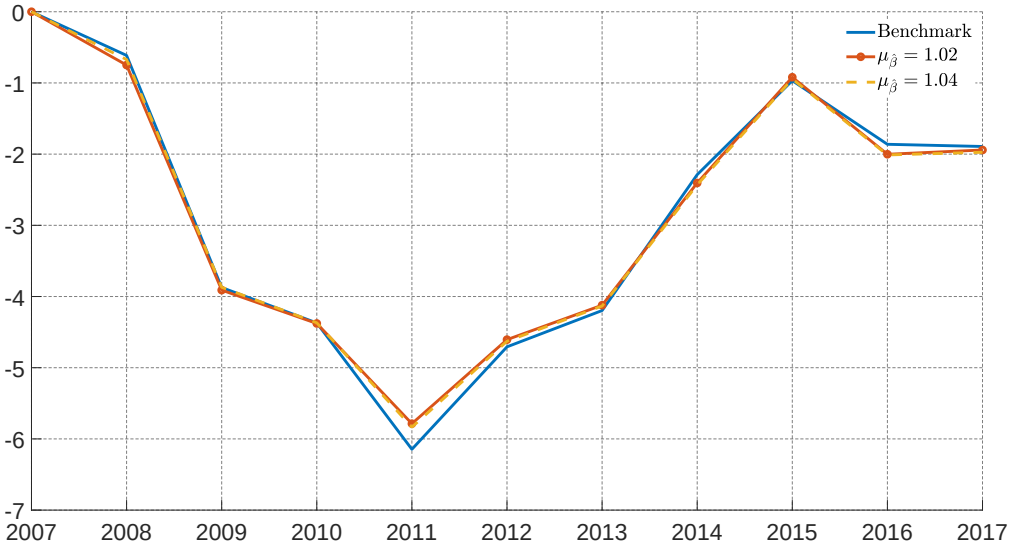


Figure A-6: Counterfactual for output, different values of $\mu_{\hat{\beta}}$

Notes: *The three lines report the posterior mean of output in the counterfactual economy with complete financial markets in percentage deviation from its 2007 value, see Online Appendix F.3 for a description of the experiment. Each line reports this counterfactual under different normalization for the average discount factor wedge. The solid line reports the benchmark ($\mu_{\hat{\beta}} = 1.00$), while the circled and dotted lines report two alternative values for $\mu_{\hat{\beta}}$.*

for two alternative values of this parameter, $\mu_{\hat{\beta}} = \{1.02, 1.04\}$. For the RA representation to fit the same data, we need to adjust the level of the discount factor β so that

$$\beta \times \mu_{\hat{\beta}} = 0.99.$$

Given this adjustment, the RA representation gives the same policy functions irrespective of the value of $\mu_{\hat{\beta}}$. The complete market counterfactual, however, features different behavior for the endogenous variables. For example, the real interest rate in the counterfactual is approximately 3% when $\mu_{\hat{\beta}} = 1.02$ and 5% when $\mu_{\hat{\beta}} = 1.04$, while it equals 1% in our benchmark. Our question is whether the different values for $\mu_{\hat{\beta}}$ also affect the business cycle properties of macroeconomic aggregates. For that purpose, we repeat the counterfactual experiment of Section 5.4 for different values of $\mu_{\hat{\beta}}$.

Figure A-6 plots the results of this experiment. The solid line reports the posterior mean of the counterfactual path of output in our benchmark ($\mu_{\hat{\beta}} = 1.00$), while the circled and dotted line report the same object for $\mu_{\hat{\beta}} = \{1.02, 1.04\}$. We can see that the three lines lie almost on top of each other: output in the counterfactual economy with complete markets falls by roughly 4% of its 2007 value during the Great Recession, irrespective of the value of $\mu_{\hat{\beta}}$. That is, while the average of the discount factor wedge affects the steady state of this economy, it does not affect much the dynamics of output around it.

To understand this result, note that changes in $\mu_{\hat{\beta}}$ in the counterfactual economy are isomorphic to changes in the rate of time preference β in the standard three equations New Keynesian model. When we express variables in log deviations from steady state, β enters the equilibrium conditions only in the forward looking component of the NK Phillips curve. Due to the degree of price stickiness that we estimate, inflation is fairly stable over time and changing β affects little the dynamics of inflation and of the other endogenous variables. This explains why different choices of $\mu_{\hat{\beta}}$ in our framework have little effects on the behavior of output in our counterfactual.