Is Household Heterogeneity Important for Business Cycles?*

Youngsoo Jang   Takeki Sunakawa   Minchul Yum

January 2019

Abstract

This paper explores how the interaction between household heterogeneity and progressive government transfers shapes aggregate labor market fluctuations. Using a static model of the extensive margin labor supply, we analytically show that greater progressivity in the transfer system leads to greater volatility of low wage workers' employment and less procyclical average labor productivity. We then build a dynamic general equilibrium model with both idiosyncratic and aggregate productivity shocks and show that household heterogeneity substantially shapes the dynamics of macroeconomic aggregates when interacted with progressive transfers. Specifically, a notable feature of the performance of our heterogeneous-agent model is its ability to reproduce moderately procyclical average labor productivity while retaining the success of the representative-agent indivisible labor model in generating a large cyclical volatility of aggregate hours relative to output. Finally, we document that among low-wage workers, (i) the individual-level probability of adjusting labor supply along the extensive margin is significantly higher, and (ii) the fall in employment rate is considerably steeper during the last six recessions, both of which support the key mechanism of our model.

Keywords: Heterogeneity, progressivity, government transfers, extensive margin labor supply, business cycles

JEL codes: E32, E24, E21

*We thank Árpád Ábrahám, Björn Brügemann, Jay Hong, Pantelis Kazakis, Aubhik Khan, Sun-Bin Kim, Tom Krebs, Claudio Michelacci, Matthias Meier, Benjamin Moll, Fabrizio Perri, Shuhei Takahashi, Michèle Tertilt, Julia Thomas and seminar participants at the Bank of Canada, European University Institute, Federal Reserve Board, Korea University, Ohio State University, Queen Mary University of London, University College London, the University of Kent, the University of Mannheim, and various conferences and workshops for helpful comments. We are grateful to Eunseong Ma and Yikai Wang for their constructive discussions on earlier versions of the paper. We gratefully acknowledge funding by the German Research Foundation (DFG) through CRC TR 224. Jang: Shanghai University of Finance and Economics, China, e-mail: jangys724@gmail.com. Sunakawa: Kobe University, Japan, e-mail: takeki.sunakawa@gmail.com. Yum: University of Mannheim, Germany, e-mail: minchul.yum@uni-mannheim.de.

1 Introduction

There has been great interest in incorporating rich micro-level heterogeneity into macroeconomic models in recent decades (see e.g., Krusell and Smith, 2006; and Heathcote, Storesletten, and Violante, 2009 for literature reviews). Clearly, it is essential to incorporate household or firm heterogeneity when studying distributional issues within a macroeconomic framework. However, it is less clear whether heterogeneity at the micro level matters for aggregate business cycle dynamics at the macro level. Although extensive studies show the importance of heterogeneity in macroeconomic aggregates and equilibrium prices in the absence of aggregate risk (e.g., Huggett, 1993; and Heathcote, 2005 among others), the recent quantitative macroeconomic literature with aggregate uncertainty has suggested that incorporating micro-level heterogeneity may have only limited impacts on the business cycle fluctuations of macroeconomic aggregates (e.g., Krusell and Smith, 1998; Thomas, 2002; Khan and Thomas, 2008; and Chang and Kim, 2007; 2014).

In this paper, we study how household heterogeneity shapes aggregate labor market fluctuations in the presence of progressive government transfers both theoretically and quantitatively. We first present a simple tractable model of the extensive margin labor supply. We analytically show that the interaction between household heterogeneity and the progressive transfer system can increase the degree to which aggregate hours vary over the business cycle and make average labor productivity less procyclical through more elastic labor supply among low-wage households. We then construct a dynamic general equilibrium incomplete-markets model with idiosyncratic and aggregate shocks, augmented with progressive government transfers. We find that our quantitative business cycle model delivers moderately procyclical average labor productivity and a large cyclical volatility of aggregate hours relative to output, both of which are known to be difficult to explain by standard real business cycle models. In particular, our result is distinct from the existing literature because our heterogeneous-agent model dampens a strong link between average labor productivity and output without relying on an additional source of exogenous shocks.[1] At the same time, our heterogeneous-agent model retains the success of the canonical representative-agent indivisible labor

[1] The existing quantitative theoretical explanations for lowering a highly procyclical labor productivity in the model rely on the introduction of additional shocks. More specifically, Benhabib et al. (1991) consider home-production technology shocks; Christiano and Eichenbaum (1992) suggest government spending shocks; Braun (1994) introduces income tax shocks; and Takahashi (2018) incorporates idiosyncratic wage uncertainty shocks into a real business cycle model.

model in generating a large volatility of aggregate hours without the assumptions of lotteries and perfect consumption insurance (Rogerson, 1988). We show that the key to our quantitative results is heterogeneity in labor supply responses at the micro level in the presence of progressive government transfers, in line with both our analytical results from the simple model and the microeconomic evidence we document using the PSID.

Our baseline model economy is based on a standard incomplete markets model with heterogeneous households that make consumption-savings and extensive-margin labor supply decisions in the presence of both idiosyncratic productivity risk and aggregate risk (Chang and Kim, 2007; 2014).[2] Our model also incorporates progressive government transfers, captured by a parsimonious yet flexible nonlinear function. We calibrate our model economy to match salient features in the micro-level data, including the degree of progressivity in the welfare programs obtained from the Survey of Income and Program Participation (SIPP) data and the persistence of idiosyncratic wage risk obtained from the Panel Study of Income Dynamics (PSID) data.

We find that our baseline model features aggregate labor market dynamics that differ considerably from its nested versions, abstracting from either government transfers (a model similar to Chang and Kim, 2007; 2014) or household heterogeneity (a model similar to Hansen, 1985). Specifically, we find significant improvements in the business cycle statistics regarding aggregate labor market fluctuations. First, our baseline model with the nonlinear government transfer schedule generates considerably lower correlations of average labor productivity with output (0.58 vs. 0.30 in the data) than its nested versions (0.95 in the absence of government transfers and 0.77 or 0.81 in the absence of household heterogeneity). At the same time, in our baseline model, the cyclical volatility of aggregate hours relative to output is 0.83, which is much closer to 0.98 in the data compared to 0.55 in the heterogeneous-agent model without government transfers.[3] It is striking that this performance is comparable to 0.85, the value obtained from its representative-agent counterpart, which is often considered to be the upper bound due to the representative agent's utility function having the lowest curvature in labor supply.

[2] This class of models in turn builds on a standard incomplete markets model without aggregate risk, pioneered by Imrohoroğlu (1988), Huggett (1993) and Aiyagari (1994).

[3] This finding is noteworthy, given that the presence of household heterogeneity seems to make it more challenging to generate a large volatility of aggregate hours, according to the recent findings in the business cycle literature. For example, Chang and Kim (2014) report that the volatility of aggregate hours relative to output is 0.58 in their model with indivisible labor.

To illustrate the key mechanism underlying our quantitative success, we conduct impulse response exercises. We find that in our baseline model, aggregate hours fall considerably more and average labor productivity does not fall deeply shortly after a persistent negative shock in total factor productivity (TFP). We also compute the impulse responses of aggregate hours at the disaggregated level. We find that labor supply is generally less elastic among households with high productivity, consistent with the analytical finding in our simple static model of the extensive margin labor supply. This pattern of heterogeneity in the labor supply (i.e., disproportionately more elastic labor supply among low-wage workers) and the resulting compositional changes following the aggregate TFP shock underlie the quantitative success of our baseline heterogeneous-agent model with government transfers. In the heterogeneous-agent model without government transfers, however, we find that the labor supply of low-wage households is remarkably inelastic due to the strong precautionary labor supply motive (Yum, 2018). This finding shows that the presence of household heterogeneity *per se* is not sufficient in an incomplete markets framework, thereby explaining why the existing heterogeneous-agent model (e.g., Chang and Kim, 2007; 2014) is unable to deliver our main quantitative results.

Finally, we use micro data from the PSID to empirically explore heterogeneity in labor supply responses.[4] We document two key empirical findings using two different approaches: the first approach uses individual-level flow data, and the second uses short-run employment rate changes shaped by aggregate factors.[5] First, we find that the individual-level probability of adjusting labor supply along the extensive margin is significantly higher among low-wage workers, and this probability tends to decrease with wages. Second, we document that during the last six recessions, the employment rate has fallen most sharply in the first wage quintile, compared to the other wage quintiles. Although the above two approaches capture labor supply adjustments over different time horizons, shaped by different forcing variables (idiosyncratic vs. aggregate factors), we find that these two findings are remarkably similar, demonstrating the robustness of our empirical result that lower-wage workers adjust the labor supply along the extensive margin more elastically. Both of these empirical findings are consistent with the pattern of heterogeneity in labor supply responses

[4] There is limited empirical evidence on heterogeneity in labor supply responses at the extensive margin. See Kydland (1984) and Juhn, Murphy and Topel (1991) for earlier evidence.

[5] We use the panel structure of the PSID, which allows us to keep track of the same individuals over time.

in our heterogeneous-agent model, thereby supporting our key model mechanism.

Our main result suggests that household heterogeneity at the micro level is important for the dynamics of macroeconomic variables. This result is broadly in line with recent papers such as Krueger, Mitman and Perri (2016) and Ahn, Kaplan, Moll, Winberry and Wolf (2017), both of which find that heterogeneity at the micro level is crucial in shaping the impact of aggregate shocks on macroeconomic variables.[6] Although the distribution of wealth plays an important role in these studies and in our study, it is important to note that Krueger et al. (2016) and Ahn et al. (2017) focus on the consumption-savings channel, whereas our paper focuses on the labor supply channel as a key mechanism through which micro-level heterogeneity matters for the business cycle fluctuations of macroeconomic aggregates, such as aggregate hours and labor productivity.

Our paper builds upon the literature that highlights the role of government transfers in affecting the precautionary behavior of low-income households. An earlier paper by Hubbard et al. (1995), for example, shows that social insurance discourages precautionary savings among low-income households. Using an incomplete-markets model without aggregate uncertainty, Yum (2018) shows that government transfers reduce the precautionary motives of employment among wealth-poor households that lack savings for self-insurance. Our results herein suggest that the presence of progressive government transfers in this class of incomplete-markets environments not only matters for the long-run employment effects of labor taxes, as studied in Yum (2018), but also has important implications for aggregate hours volatility over the business cycle. Our paper further contributes to this literature by providing the analytical results on the interaction between household heterogeneity and the progressivity of government transfers, as well as the quantitative results on how the dynamics of average labor productivity over the business cycle can be affected by this mechanism.

Our quantitative business cycle model is based on Chang and Kim (2007; 2014), but it differs from theirs in at least two important respects. First, as highlighted above, we bring the institutional feature of progressive government transfers, as observed in the micro-level data (SIPP), into the model.[7] Second, note that the estimation of the idiosyncratic productivity (wage) process for

[6] See also Kim (2018).

[7] Chang, Kim, and Schorfheide (2013) consider a version of the model in Chang and Kim (2007) with flat lump-sum transfers. However, given the different focus of their paper, they report a limited number of standard business cycle

the model with the extensive margin labor supply is not trivial because wages are observable in the data only for those who choose to work. Chang and Kim (2007; 2014) address this selection problem outside the model by applying the Heckman (1979) correction. In contrast, we address the selection problem and potential temporal aggregation bias (quarterly model vs. annual micro-level data) using the model simulation directly. Specifically, we use the simulated data where selection is endogenously addressed within the model and then perform temporal aggregation using the simulated quarterly data to obtain the simulated annual data. Moreover, our calibration targets the persistence of idiosyncratic wage risk estimated following Heathcote, Storesletten, and Violante (2010). As a result, our calibration strategy leads to a considerably high persistence of idiosyncratic productivity shocks, which interacts with the presence of progressive government transfers in improving the performance of the incomplete-markets business cycle model.[8]

The paper is organized as follows. Section 2 presents analytic results on the key mechanism of this paper. Section 3 describes the model environment of the equilibrium business cycle models, defines equilibrium, and discusses the numerical solution methods. In Section 4, we describe how the parameters are calibrated and show the properties of the quantitative models in stationary equilibrium. Section 5 presents the main quantitative results from the calibrated models. Section 6 presents empirical evidence on heterogeneity in labor supply using the panel structure of the PSID. Section 7 concludes the paper.

2 Interplay between household heterogeneity and progressive transfers

In this section, we present a simple model of the extensive margin labor supply to illustrate the key mechanism through which the interplay between household heterogeneity and progressive transfers influences aggregate labor market fluctuations. For analytical tractability and a clear illustration,

statistics, which are the main focus of our paper.

[8]Specifically, the persistence estimate of the idiosyncratic productivity shocks at the annual frequency ranges from 0.91 to 0.93 in our paper, whereas it is approximately 0.75 in Chang and Kim (2007). Our calibrated value is quite similar to the value used in Chang, Kim, Kwon and Rogerson (2019). In Section 2, we discuss how this persistence matters for our results.

we consider a static economic environment by taking the distribution of wealth as given.[9]

There is a continuum of agents in the unit interval. We assume that there are two types of productivity. That is, individual productivity x_i can be either low or high: $x_i \in \{x_l, x_h\}$. The mass of each type is denoted by π_l and π_h satisfying $\pi_l + \pi_h = 1$. A subscript $i \in \{l, h\}$ denotes the type of the agent throughout this section. Agents are allowed to differ also in their level of asset holdings, a. Because our focus is on the extensive margin, the agent can choose to either work full time or not at all: $n_i \in \{0, 1\}$.[10]

The decision problem of each type is given by

$$\max_{c_i \geq 0, n_i \in \{0,1\}} \{\log c_i - b n_i\}$$

subject to

$$c_i \leq z x_i n_i + a + T_i, \quad i = l, h$$

where c denotes consumption, n is the employment choice, a is the level of assets, and $b > 0$ captures the disutility of work. We use z to denote aggregate productivity. Finally, we introduce a productivity-dependent public insurance scheme $T_i \geq 0$. We assume that T_l is greater than T_h, implying that it is progressive.[11]

The above maximization problem describes the optimal decisions of an individual. Specifically, comparing the utility conditional on working to not working, the agent chooses to work if

$$\log (z x_i + T_i + a) - b \geq \log (T_i + a).$$

Note that this can be equivalently written as

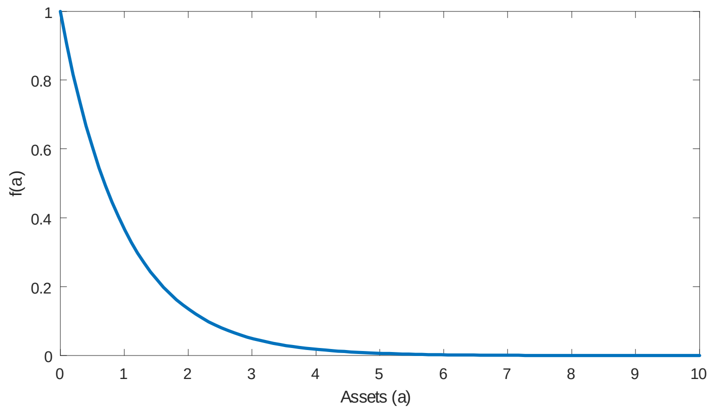
$$b \leq \log \left(\frac{z x_i + T_i + a}{T_i + a} \right) = \log \left(1 + \frac{z x_i}{T_i + a} \right),$$

[9] The quantitative model in the following sections relaxes this assumption and the distribution of wealth is endogenously determined by the household consumption-savings decision.

[10] Our analytical framework in this section builds on the theoretical framework in Doepke and Tertilt (2016). Since the focus of the analysis is different, their model is based on two gender types and continuous preference heterogeneity whereas our model is based on two productivity types and continuous asset heterogeneity. Moreover, our results cover not only the labor supply elasticity but also average labor productivity.

[11] In this section, we assume that transfers depend on productivity for analytical tractability. In the following sections, our quantitative exercises relax this assumption and use income-dependent transfers instead.

Figure 1: Probability density function of assets in the static model



or

$$a \leq zx_i - T_i$$

where we assume that the constant b is equal to $\log(2) > 0$ without loss of generality. This decision rule shows that the agent is more likely to choose to work if the aggregate condition z or individual productivity x is higher. Also, note that the agent is less likely to choose to work if the size of transfers is higher.

In our model of the extensive margin labor supply, aggregate employment is shaped by both the decision rule and the distribution. Let $F_i(a)$ be the conditional (differentiable) distribution function of assets with its marginal density being $f_i(a) = F_i'(a)$. Specifically, we use the exponential function for our following results: for $a \geq 0$,

$$F_i(a) = 1 - \exp(-a),$$
$$f_i(a) = F_i'(a) = \exp(-a).$$

This density function has the mode at 0 and is strictly decreasing in a, as shown in Figure 1. In addition, it generates a long right tail of asset distribution, as in the data.[12]

Given the density function and the work decision rule, the fraction of agents working (i.e., the

[12] We note that our results are more general in that it is possible to derive the same theoretical results with alternative tractable probability distributions featuring these properties.

employment rate) for each type is given by

$$N_i = F(\tilde{a}_i) = 1 - \exp(-\tilde{a}_i)$$

where

$$\tilde{a}_i = zx_i - T_i.$$

In other words, the employment rate of the type i, N_i, is the integral of the type i agents whose asset level is lower than the threshold level $\tilde{a}_i$. We now present some theoretical results based on this model. All proofs are provided in Appendix A.

Proposition 1 *Let ε_i be the labor supply elasticity of the type i.*

$$\varepsilon_i \equiv \frac{\partial N_i}{\partial z}\frac{z}{N_i}.$$

Assume $T_i = 0$. The labor supply elasticity of the low type is greater than that of the high type. That is, $\varepsilon_l > \varepsilon_h$.

This shows that our model of the extensive margin delivers heterogeneity in the labor supply elasticity. Note that the threshold asset level of employment for the low-type agents is lower than that for the high-type agents: $\tilde{a}_l < \tilde{a}_h$. Since there are more marginal households around $\tilde{a}_l$, as shown in Figure 1, the same change in z, which shifts the employment thresholds to the same degree, has a larger impact on the employment rate of the low type.

We now consider the role of government transfers and how they interact with heterogeneity. To simplify the algebra, we impose symmetry. Specifically, we assume that $\pi_l = \pi_h = 0.5$. In addition, $x_h = 1 + \lambda$ and $x_l = 1 - \lambda$ where $\lambda \in [0, 1]$ measures the cross-sectional dispersion of productivity.

To study the effect of progressivity in the transfer schedule, T_i is assumed to be determined by

$$T_l = T(1 + \omega\lambda)$$
$$T_h = T(1 - \omega\lambda)$$

where $T \in [0, z]$ is the scale of transfers, and $\omega \in [0, \frac{1}{\lambda}]$ captures progressivity. A changes in progressivity ω does not affect the aggregate amount of transfers.[13]

Given the above assumptions, we can derive the type-specific employment rate:

$$N_l = 1 - \exp(-\tilde{a}_l)$$

$$N_h = 1 - \exp(-\tilde{a}_h)$$

where $\tilde{a}_l = \max\{0, z(1-\lambda) - T - T\omega\lambda\}$ and $\tilde{a}_h = \max\{0, z(1+\lambda) - T + T\omega\lambda\}$.

Proposition 2 *Greater progressivity of transfers increases the labor supply elasticity of the low-type agents, while it decreases the labor supply elasticity of the high-type agents. That is,* $\frac{\partial \varepsilon_l}{\partial \omega} > 0$ *and* $\frac{\partial \varepsilon_h}{\partial \omega} < 0$.

Intuitively, greater progressivity (or a higher ω) shifts $\tilde{a}_l$ to the left where the distribution of assets is denser. There, the same change in z would influence more agents' employment decisions, thereby leading to a greater elasticity of the low type. In contrast, greater progressivity shifts $\tilde{a}_h$ to the right around which the distribution of assets is scarcer. A fewer number of agents around the new employment threshold implies that the elasticity of the high type agents would become smaller.

Proposition 3 *Let* N *denote the aggregate employment rate:* $N = \pi_l N_l + \pi_h N_h$. *Let* ε *be the aggregate labor supply elasticity:*

$$\varepsilon \equiv \frac{\partial N}{\partial z} \frac{z}{N}.$$

The aggregate labor supply elasticity is higher with greater progressivity. That is, $\frac{\partial \varepsilon}{\partial \omega} > 0$.

This result is straightforward given Propositions 1 and 2 and the fact that $f(a)$ is more concentrated as a becomes lower. This result underlies one of our main findings in Section 5 showing that incorporating progressive government transfers allows the quantitative business cycle model with household heterogeneity to generate a large volatility of aggregate hours over the business cycle.

Next, we consider the implications for the cyclicality of average labor productivity. Define

[13] Note that $\sum \pi_i T_i = \pi_l T(1+\omega\lambda) + \pi_h T(1-\omega\lambda) = T$.

average labor productivity as

$$\chi \equiv \frac{\sum_{j\in\{l,h\}} \pi_i \left(z x_i N_i\right)}{\sum_{j\in\{l,h\}} \pi_i N_i} = z \frac{\sum_{j\in\{l,h\}} \pi_i \left(x_i N_i\right)}{\sum_{j\in\{l,h\}} \pi_i N_i}.$$

Letting $\chi = z\chi_0$, we can see that aggregate productivity z would directly make average labor productivity procyclical, as in real business cycle models. In addition, the second term χ_0 captures the effects of worker composition on average labor productivity. Note that χ_0 also depends indirectly on z through type-specific employment responses. The following two propositions focus on the second term.

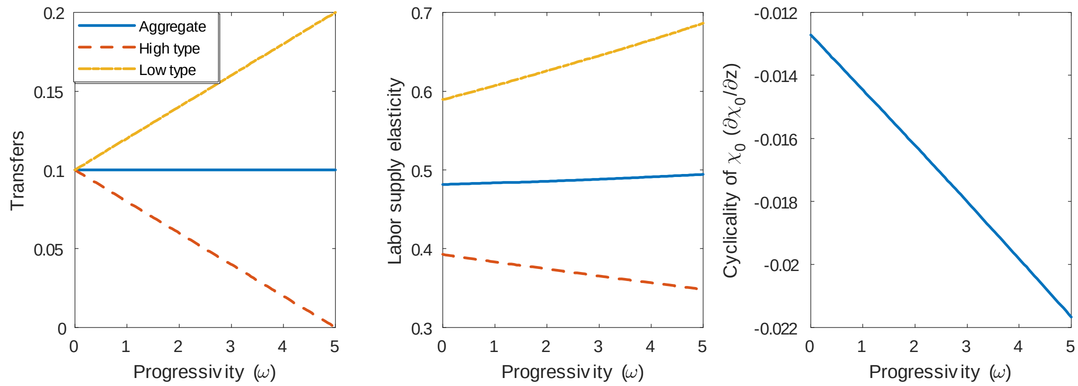
Proposition 4 *The effect of z on average labor productivity through worker composition effects is negative. That is, $\frac{\partial \chi_0}{\partial z} < 0$.*

Proposition 5 *Average labor productivity becomes less procyclical with greater progressivity. That is, $\frac{\partial}{\partial \omega}\left(\frac{\partial \chi_0}{\partial z}\right) < 0$.*

Proposition 5 tells us that progressivity in the transfer schedule shapes the cyclicality of average labor productivity through worker composition effects. Specifically, it implies that greater progressivity would make average labor productivity *less* procyclical, as illustrated in a numerical example in the right panel of Figure 2. The intuition behind this result is in fact related to Propositions 1 and 2. Note that the positive impact of progressivity on aggregate labor supply responsiveness in Proposition 3 (as depicted in the left panel of Figure 2) is driven by the low type having a stronger labor supply response to a change in z. If a fall in z (e.g., in a recession) generates a large fall in the labor supply of the low type (especially relative to the high type), it would then tend to raise average labor productivity, while output falls. This force dampens the tight positive link between z and average labor productivity.

Before we move on, it is worth discussing our assumption that $F_l(a) = F_h(a)$. In other words, the conditional distribution of assets for each type is assumed to be identical. In fact, it should be noted that this assumption is quite conservative. For example, if we assume that $F_i(a) = 1 - \exp(-\frac{1}{\mu_i}a)$ and $\mu_l < \mu_h$ (as is consistent with the data), the distribution of assets for the low type would become more packed around the threshold asset level, which in turn would strengthen the above results (due

Figure 2: Impact of progressivity on aggregate labor market



Note: A numerical example of Propositions 1-5, based on $b = 2$, $\lambda = 0.2$, $T = 0.1$, and $z = 1.53$ to match the aggregate employment rate of 75% when $\omega = 0$.

to the even more elastic labor supply of the low-type). Furthermore, this highlights the importance of the persistence of idiosyncratic shocks in a dynamic environment where the distribution of assets is endogenously determined because higher persistence would enlarge the difference of the mean of assets across idiosyncratic productivity types.[14]

In the next sections, we imbed these key insights into standard equilibrium business cycle models with more realistic household heterogeneity to explore how this mechanism can alter the model-implied dynamics of macroeconomic aggregates.

3 Quantitative business cycle models

In this section, we describe the economic environment of the quantitative business cycle models studied in this paper. We consider four model specifications. The first is the baseline model with household heterogeneity and progressive government transfers. The other alternative specifications are considered to illustrate the importance of the interplay between household heterogeneity and government transfers.

[14] As an extreme example, consider the case where the idiosyncratic shock has zero persistence (i.i.d.). In this case, the equilibrium asset distribution for each type would become identical, as in this section.

3.1 Heterogeneous-agent models

In this subsection, we introduce the first two model specifications with heterogeneous households. The first is the baseline model with government transfers, denoted as Model (HA-T). The second, denoted as Model (HA-N), is simply a nested specification of the baseline model by shutting down government transfers. This model roughly corresponds to a standard incomplete-markets real business cycle model with household heterogeneity and endogenous labor supply at the extensive margin (Chang and Kim, 2007; 2014). In other words, the baseline model economy extends Chang and Kim (2007, 2014) by incorporating labor taxes and progressive government transfers.

Households:

The model economy is populated by a continuum of infinitely-lived households. It is convenient to describe the infinitely-lived household's decision problem recursively. At the beginning of each period, households are distinguished by their asset holdings a and productivity x_i. We assume that x_i takes a finite number of values N_x and follows a Markov chain with transition probabilities π^x_{ij} from the state i to the state j. In addition to the individual state variables, a and x_i, there are aggregate state variables, including the distribution of households $\mu(a, x_i)$ over a and x_i and aggregate total factor productivity shocks z_k. We also assume that z_k takes a finite number of values N_z following a Markov chain with transition probabilities π^z_{kl} from the state k to the state l. We assume that the Markov processes for individual productivity x and aggregate productivity z capture the following continuous AR(1) processes in logs.

$$\log x' = \rho_x \log x + \varepsilon'_x \tag{1}$$

$$\log z' = \rho_z \log z + \varepsilon'_z \tag{2}$$

where $\varepsilon_x \sim N(0, \sigma_x^2)$ and $\varepsilon_z \sim N(0, \sigma_z^2)$. A variable with a prime denotes its value in the next period. Finally, we assume competitive markets; in other words, households take as given the wage rate per efficiency unit of labor $w(\mu, z_k)$ and the real interest rate $r(\mu, z_k)$, both of which depend on the aggregate state variables. Households take as given government policies.

The dynamic decision problem of households can be written as the following functional equation:

$$V(a, x_i, \mu, z_k) = \max \left\{ V^E(a, x_i, \mu, z_k), V^N(a, x_i, \mu, z_k) \right\}$$

where

$$V^E(a, x_i, \mu, z_k) = \max_{\substack{a' > \underline{a}, \\ c \geq 0}} \left\{ \log c - B\bar{n} + \beta \sum_{j=1}^{N_x} \pi_{ij}^x \sum_{l=1}^{N_z} \pi_{kl}^z V(a', x_j', \mu', z_l') \right\} \tag{3}$$

$$\text{subject to} \quad c + a' \leq (1-\tau) w(\mu, z_k) x_i \bar{n} + (1 + r(\mu, z_k)) a + T(w(\mu, z_k) x_i \bar{n} + r(\mu, z_k) a)$$

$$\mu' = \Gamma(\mu, z_k).$$

and

$$V^N(a, x_i, \mu, z_k) = \max_{\substack{a' > \underline{a}, \\ c \geq 0}} \left\{ \log c + \beta \sum_{j=1}^{N_x} \pi_{ij}^x \sum_{l=1}^{N_z} \pi_{kl}^z V(a', x_j', \mu', z_l') \right\} \tag{4}$$

$$\text{subject to} \quad c + a' \leq (1 + r(\mu, z_k)) a + T(r(\mu, z_k) a)$$

$$\mu' = \Gamma(\mu, z_k).$$

Households maximize utility by choosing optimal consumption c, asset holdings in the next period a', and labor supply n. The labor supply decision is assumed to be discrete $n \in \{0, \bar{n}\}$. The disutility of work is captured by $B > 0$. Households understand that the expected future value, discounted by a discount factor β, is affected by stochastic processes for individual productivity x' and aggregate productivity z' as well as the whole distribution μ'. The evolution of μ is governed by the law of motion $\mu' = \Gamma(\mu, z_k)$. The budget constraint states that the sum of current consumption c and asset demands for the next period a' should be less than or equal to the sum of net-of-tax earnings $(1-\tau)w(\mu, z_k)x_i n$, current asset holdings and capital income $(1 + r(\mu, z_k))a$, and government transfers $T(\cdot)$. We allow transfers to depend on income, and we elaborate on it below. Households face a borrowing limit $\underline{a} = 0$.[15]

[15] We have considered a version of the model that incorporates a moderate non-zero borrowing limit. The main results found in this paper are barely affected by this.

Government:

There is a government that taxes labor earnings at a fixed rate of τ. The government uses the collected tax revenue to finance transfers to households. The remaining tax revenue is spent as government spending, which is not valued by households.

Following Krusell and Rios-Rull (1999), we assume that transfers T consist of two components (i.e., $T = T_1 + T_2$). The first component T_1 is given to all households equally whereas the second component T_2 captures the income security aspect of transfers. In the U.S., there are various means-tested programs such as food stamps, the Supplemental Nutrition Assistance Program and the Temporary Assistance for Needy Families (formerly the Aid to Families with Dependent Children). As shown in Section 4, these programs lead to the observation that the amount of transfers is negatively associated with income. We assume that T_2 depends on total household income m to replicate the progressivity observed in the U.S. data using the following functional form (Yum, 2018):[16]

$$T_2(m) = \omega_s(1+m)^{-\omega_p}. \tag{5}$$

This parametric assumption adds two parameters. First, $\omega_s \geq 0$ is a scale parameter, which determines the overall size of progressive part of government transfers (i.e., T_2). The next parameter $\omega_p \geq 0$ governs the degree of progressivity: a higher ω_p would make T_2 decrease faster with income. Note that $\omega_p = 0$ would imply that the transfer schedule is independent of income, which is commonly used in the literature.

Firm:

Aggregate output Y is produced by a representative firm. The firm maximizes its profit

$$\max_{K,L} \{z_k F(K,L) - (r(\mu, z_k) + \delta)K - w(\mu, z_k)L\} \tag{6}$$

where $F(K,L)$ captures a standard neoclassical production technology in which K denotes aggregate capital, L denotes aggregate efficiency units of labor inputs, and δ is the capital depreciation rate. As is standard in the literature, we assume that the aggregate production function follows a

[16] Alternatively, we have also considered a verion of the model with a productivity-dependent transfer schedule (e.g., Oh and Reis, 2012), and found that the main results on business cycles are quite similar to the current version.

Cobb-Douglas function with constant returns to scale:

$$F(K, L) = K^{\alpha} L^{1-\alpha}. \tag{7}$$

The first-order conditions for K and L give

$$r(\mu, z_k) = z_k F_1(K, L) - \delta, \tag{8}$$

$$w(\mu, z_k) = z_k F_2(K, L). \tag{9}$$

Equilibrium:

A recursive competitive equilibrium is a collection of factor prices $r(\mu, z_k), w(\mu, z_k)$, household decision rules $g_a(a, x_i, \mu, z_k)$, $g_n(a, x_i, \mu, z_k)$, government policy variables $\tau, G, T(\cdot)$, a value function $V(a, x_i, \mu, z_k)$, a distribution of households $\mu(a, x_i)$ over the state space, the aggregate capital and labor $K(\mu, z_k), L(\mu, z_k)$ and the aggregate law of motion $\Gamma(\mu, z_k)$ such that

1. Given factor prices $r(\mu, z_k), w(\mu, z_k)$ and government policy $\tau, G, T(\cdot)$, the value function $V(a, x_i, \mu, z_k)$ solves the household's decision problems defined above, and the associated household decision rules are

$$a'^{*} = g_a(a, x_i, \mu, z_k) \tag{10}$$

$$n^{*} = g_n(a, x_i, \mu, z_k). \tag{11}$$

2. Given factor prices $r(\mu, z_k), w(\mu, z_k)$, the firm optimally chooses $K(\mu, z_k)$ and $L(\mu, z_k)$ following (8) and (9).

3. Markets clear:

$$K(\mu, z_k) = \sum_{i=1}^{N_x} \int_a a d\mu \tag{12}$$

$$L(\mu, z_k) = \sum_{i=1}^{N_x} \int_a x_i g_n(a, x_i, \mu, z_k) d\mu. \tag{13}$$

4. Government balances its budget:

$$G + \sum_{i=1}^{N_x} \int_a T(m) d\mu = \tau w L(\mu, z_k).$$

5. The law of motion for the distribution of households over the state space $\mu' = \Gamma(\mu, z_k)$ is consistent with individual decision rules and the stochastic processes governing x_i and z_k.

3.2 Representative-agent models

In addition to the heterogeneous-agent model specifications, i.e., Models (HA-T) and (HA-N), we consider two additional specifications of the representative-agent model. First, Model (RA-T) shuts down household heterogeneity while maintaining the fiscal environment including taxes and transfers, as in Model (HA-T). Given the indivisible labor assumption, our representative-agent version of the model is essentially the business cycle model studied in Hansen (1985) augmented with tax and transfers. Next, Model (RA-N) shuts down both household heterogeneity and government transfers. We consider the decentralized competitive equilibrium given distortionary labor taxation.

Representative-agent model environment:

At the beginning of each period, the stand-in household has the current period's assets k. The aggregate state variables are the aggregate capital K and the aggregate productivity z_k. The aggregate productivity follows the same stochastic process as in the baseline model. Taking the wage rate $w(K, z_k)$ and the real interest rate $r(K, z_k)$, as well as the aggregate law of motion $\Gamma(K, z_k)$ as given, the dynamic decision problem of the representative household can be written as the following functional equation:

$$V(k, K, z_k) = \max_{\substack{k' \geq 0, c \geq 0 \\ n \in [0,1]}} \left\{ \log c - Bn + \beta \sum_{l=1}^{N_z} \pi^z_{kl} V(k', K', z'_l) \right\}$$

$$\text{subject to} \quad c + k' \leq (1 - \tau) w(K, z_k) n + (1 + r(K, z_k)) k + T_1$$
$$K' = \Gamma(K, z_k)$$

The household maximize utility by choosing optimal consumption c, the next period's capital k' and labor supply n. The stand-in household's utility is linear in employment n due to the aggregation theory in Rogerson (1988). The budget constraint states that the sum of consumption c and the next period's capital k' should be less than or equal to the sum of net-of-tax labor income $(1 - \tau) w(K, z_k) n$, current capital k, capital income $r(K, z_k) k$ and government transfers T.

Government collects taxes on labor earnings $\tau w n$ to finance transfers T_1 and government expenditure G. We keep the same assumptions on the firm side, as in the heterogeneous-agent models. The resulting first-order conditions for K and L are the same as those in (8) and (9).

Equilibrium:

A recursive competitive equilibrium is a collection of factor prices $r(K, z_k)$, $w(K, z_k)$, household decision rules $g_k(k, K, z_k)$, $g_n(k, K, z_k)$, government policy variables τ, G, T, the household value function $V(k, K, z_k)$, the aggregate labor $L(K, z_k)$ and the aggregate law of motion for aggregate capital $\Gamma(K, z_k)$ such that

1. Given factor prices $r(K, z_k)$, $w(K, z_k)$ and government policy τ, G, T, the value function $V(k, K, z)$ solves the household's decision problem, and the associated decision rules are

$$k'^* = g_k(k, K, z_k)$$
$$n^* = g_n(k, K, z_k).$$

2. Prices $r(K, z_k)$, $w(K, z_k)$ are competitively determined following (8) and (9).

3. Government balances its budget:

$$G + T = \tau w(K, z_k)L(K, z_k).$$

4. Consistency is satisfied: for all K,

$$K' = \Gamma(K, z_k) = g_k(K, K, z_k)$$
$$L(K, z_k) = g_n(K, K, z_k).$$

3.3 Solution method

Our heterogeneous-agent models (i.e., Model (HA-T) and Model (HA-N)) cannot be solved analytically, are thus solved numerically. Several key features make the numerical solution method nontrivial. First, key decision variables in our model are a discrete employment choice and a consumption-savings choice in the presence of a borrowing constraint. Therefore, our solution method is based on the nonlinear method (i.e., the value function iteration) applied to the recursive representation of the problem described above. Second, the aggregate law of motion and state variables involve an infinite-dimensional object: the distribution μ. Therefore, we solve the model by approximating the distribution of wealth by the mean of the distribution (Krusell and Smith, 1998). In addition, since market-clearing is nontrivial in our model with endogenous labor, our solution method incorporates a step to find market-clearing prices in each period when simulating the model.

We describe the solution method briefly.[17] Following Krusell and Smith (1998), we assume that households use a smaller object that approximates the infinite-dimensional distribution when they forecast the future state variables to make current decisions. More precisely, we approximate $\mu(a, x_i)$ by its mean of the asset distribution $K = \int_a \sum_{i=1}^{N_x} a d\mu$. Also, the next period's aggregate capital K', real wage rate w and real interest rate r are assumed to be functions of (K, z) instead of (μ, z). We impose the parametric assumptions to approximate the aggregate law of motion

[17] See Appendix F for more details.

$K' = \Gamma(K, z)$ and $w = w(K, z)$ following

$$\hat{K}' = \hat{\Gamma}(K, z) = \exp\left(a_0 + a_1 \log K + a_2 \log z\right) \tag{14}$$

$$\hat{w} = \hat{w}(K, z) = \exp\left(b_0 + b_1 \log K + b_2 \log z\right), \tag{15}$$

as in Chang and Kim (2007, 2014) and Takahashi (2014, 2018). Based on these forecasting rules, households obtain a forecasted $\hat{r}$ implied by the first-order conditions of firm's profit maximization problem.

Given the above forecasting rules, the model is solved in the two steps. First, we solve for the individual policy functions given the forecasting rules using the value function iterations (*the inner loop*). Then, we update the forecasting rules by simulating the economy using the individual policy functions (*the outer loop*). As noted above, it is important to note that, since our model environment with endogenous labor supply involves non-trivial factor market clearing, we incorporate a step to find market-clearing factor prices in the outer loop (Chang and Kim, 2014; Takahashi, 2014). We repeat this procedure until the coefficients in the forecasting rules converge.

It is more straightforward to solve the representative-agent version of the model. Due to the distortionary tax, we solve the decentralized competitive equilibrium. For the purpose of comparison, we keep the same assumptions on the discretization of the aggregate productivity process as in the heterogeneous-agent model. The steady-state equilibrium can be obtained analytically. For solutions with aggregate uncertainty, we use the policy function iteration method.

4 Calibration and model properties in steady state

The model is calibrated to U.S. data. A period in the model is a quarter, as is standard in the business cycle literature. There are two sets of parameters. The first set of parameters is calibrated externally, in line with the business cycle literature. These parameter values are commonly set in all model specifications. The second set of parameters is calibrated to match the same number of relevant target statistics.

We begin with describing the first set of externally-calibrated parameters. Most of these parameters are commonly used in the real business cycle literature. The capital share, α, is chosen to

Table 1: Parameter values chosen internally

	Model				
	(HA-T)	(HA-N)	(RA-T)	(RA-N)	Description
$B =$	.623	.880	.918	1.09	Disutility of work
$\beta =$	.987	.985	.990	.990	Subject discount factor
$\rho_x =$	.983	.976	-	-	Persistence of $\ln x$
$\sigma_x =$	.103	.132	-	-	S.D. of innovations to $\ln x$
$T_1 =$	.103	-	.295	-	Overall transfer size
$\omega_s =$	.091	-	-	-	Progressive transfer scale
$\omega_p =$	3.34	-	-	-	Progressivity of transfers

Note: Model (HA-T) is the baseline specification: a heterogeneous-agent model with government transfers. Model (HA-N) shuts down government transfers but keeps household heterogeneity. Model (RA-T) abstracts from household heterogeneity but keeps government transfers. Model (RA-N) shuts down both heterogeneity and government transfers. All model specifications have the same labor taxation.

be consistent with the capital share of 0.36. The quarterly depreciate rate, δ, is 2.5 percent. In our model specifications with a binary labor supply choice, the level of hours worked, $\bar{n}$, can be arbitrarily set since it simply determines the scale of the calibrated disutility parameter B. We set it to 1/3, implying that working individuals spend a third of their time endowment on working. The labor income tax, τ, is set to 27.9 percent (Trabandt and Uhlig ,2011; and Yum, 2018), based on the method proposed by Mendoza et al. (1994). Finally, it is worth noting that the main goal of this paper is to study how heterogeneity interacts with progressive transfers in the presence of aggregate shocks. As a first step, we consider total factor productivity shocks (Kydland and Prescott, 1982) as aggregate risk, and employ standard values of $\rho_z = 0.95$ and $\sigma_z = 0.007$ (Cooley and Prescott, 1995), which are also used by recent related papers such as Chang and Kim (2007, 2014) and Takahashi (2018).[18]

The second set of parameters is jointly calibrated in each specification of the model. As shown in Table 1, there are seven parameters in Model (HA-T), and four parameters in Model (HA-N) that shuts down government transfers. Unlike the heterogeneous-agent models that require simulation to calibrate these parameters, the representative-agent models are easy to calibrate

[18] An interesting exercise to be followed in the future is to investigate how our results would carry over in the presence of other types of aggregate shocks on top of the standard TFP shocks. The estimation of multiple aggregate shocks within a model with heterogeneous agents is an important yet difficult task at this stage due to the computational costs.

using the analytical optimality conditions (see Appendix E). Our discussion herein focuses on the heterogeneous-agent model specifications. The parameter values reported in Table 1 are the calibrated values by matching the same number of target statistics summarized in Table 2.

We now explain how each parameter is linked to the target statistics. The first parameter is B, which captures the disutility of work. The most relevant target moment is the employment rate of 75.0 percent in our SIPP samples. The next parameter β captures the discount factor of households. As is standard in the literature, it is targeted to match the quarterly interest rate of 1 percent.

The next two parameters, ρ_x and σ_x, govern the dynamics of idiosyncratic labor productivity. Note that there are two issues that are worth highlighting regarding how to calibrate these parameters. First, there is a discrepancy in time frequency between the model and the data: the model period is a quarter, while the wage data that are widely used to estimate wage or earnings processes in the literature are at the annual frequency. Thus, a simple transformation of an annual persistence estimate to a quarterly value may be subject to a nontrivial temporal aggregation bias because labor supply is endogenous at a higher frequency in the model. Furthermore, there is a well-known selection issue because we only observe wages if households choose to work. To address both issues, we first estimate the persistence of idiosyncratic wage risk using the PSID following a standard method in the literature (e.g., Heathcote et al. 2010). The estimation result shows that the persistence of wages at the annual frequency is 0.953, in line with the previous estimates in the literature. Then, we calibrate the model so that the persistence of annual wages that are constructed by the simulated quarterly data from the model matches 0.953.[19] Next, the standard deviation of innovations to the AR(1) process, σ_x, in (1) is calibrated to match the overall dispersion of annual earnings that are also directly constructed by the simulated quartely data in the model. This calibration strategy makes sure that the two heterogeneous-agent models have the same degree of the observed earnings inequality in U.S. data, measured by the standard deviation of log earnings at the annual frequency (0.623).

[19] More precisely, we simulate the model and construct annual wages by dividing annual earnings by annual hours worked through explicit temporal aggregation. Note that although labor supply is a binary choice at the quarterly frequency, there are richer variations in the *annual* hours worked driven by the number of quarters worked, affected by both idiosyncratic and aggregate shocks. Erosa, Fuster and Kambourov (2016) highlight a similar point in a stationary environment in the absence of business cycles.

Table 2: Target statistics in the data and in the model

Target	Data	Model (HA-T)	(HA-N)	(RA-T)	(RA-N)
Employment rate	.750	.750	.750	.750	.750
Real interest rate	.010	.010	.010	.010	.010
Persistence of annual worker wages	.953	.953	.953	-	-
S.D. of log annual worker earnings	.623	.622	.623	-	-
Ratio of $T_1 + T_2$ to output	.102	.102	-	.102	-
Ratio of T_2 to output	.020	.020	-	-	-
$E(T_2 \| \text{1st income quintile})/E(T_2)$	3.01	3.01	-	-	-

Note: See Table 1 for the description of the model specifications.

The last three parameters, T_1, ω_s and ω_p, govern statistics regarding transfers. The first target statistic regarding transfers is set as an aggregate transfers-output ratio of 10.2%, which is obtained as the time-series average of the ratio of transfers to output over the years 1961-2016 according to the BEA data. Recall that the two components of transfers can be distinguished in Model (HA-T): the parameter T_1 determines the size of flat government transfers and ω_s determines the scale of progressive transfers (T_2). Therefore, Model (HA-T) has an additional target statistic regarding the size of progressive transfers. As for this, we compute the average of income-security related government expenditures on social benefits (Table 3.12) over the years 1961-2016 (2.0% of output) in the BEA data.[20] Since Model (RA-T) lacks heterogeneity, T_2 is irrelevant. Next, note that ω_p shapes the degree of progressivity in government transfers. Our calibration strategy is to let the model to replicate an empirically reasonable degree of trasnfer progressivity through ω_p. For this purpose, we measure the degree of progressivity in the U.S. transfer programs using the SIPP data. We construct a broad measure of government transfers, including means-tested programs and social insurance (as detailed in Appendix C). Since these welfare-related programs are highly relevant for the poor households, we choose the ratio of the average means-tested transfers received by the first income quintile to its unconditional mean (3.01) as a target statistic, as in Yum (2018).

Table 2 shows that all model specifications do a good job of matching the target statistics.

[20] We select the components to be consistent with our measure of transfers in the SIPP data, as describe below. Our classification of transfers are similar to Krusell and Rios-Rull (1999). See Appendix B for details.

Table 3: Characteristics of wealth distribution

Unit: %	Wealth quintile				
	1st	2nd	3rd	4th	5th
Share of wealth					
U.S. Data	-1.6	1.7	7.8	19.0	73.9
Model (HA-T)	0.2	1.2	7.8	20.5	70.3
Model (HA-N)	0.0	1.5	9.6	23.4	65.5
Employment rate					
U.S. Data	68.6	76.2	77.4	78.1	74.5
Model (HA-T)	72.2	87.3	76.5	76.3	62.7
Model (HA-N)	100.0	88.6	74.7	59.5	52.0

Note: The source of U.S. data is the Survey of Income and Program Participation 2001. See Table 1 for the description of the model specifications.

This does not necessarily mean that the model does a good job of accounting for other relevant statistics. We thus present (non-targeted) distributional aspects of the model economy in steady state. First, Table 3 summarizes the share of wealth, employment rates by wealth quintile. Overall, both heterogeneous-agent model specifications do a good job of accounting for the share of wealth by wealth quintile. A closer look reveals that Model (HA-T) does a better job of accounting for the wealth concentration at the top of the wealth distribution. Specifically, the relative shares of the fourth and fifth quintiles are noticeably closer to the data (19.0% and 73.9%, respectively) in Model (HA-T) (20.5% and 70.3%, respectively) compared to Model (HA-N) (23.4% and 65.5%, respectively). Because that the presence of government transfers reduces households' incentive to save (Hubbard, Skinner and Zeldes, 1995), the relative share of wealth by households in the top wealth quintile that receive little transfers becomes larger in Model (HA-T) with progressive government transfers.

When we look at the employment rate by wealth quintile, it is clear that Model (HA-T) does a significantly better job of accounting for the cross-sectional employment-wealth relationship. In the U.S., the employment rate of the first wealth quintile is relatively low (68.6%), compared to that of the second quintile (76.2%), and then it declines with wealth. This weakly inverse-U shape of the employment rates across wealth quintiles in the data are well captured in Model (HA-T). On the other hand, Model (HA-N) predicts that employment falls sharply with wealth, consistent

Table 4: Progressivity of income-security transfers

Unit: %	Income quintile				
	1st	2nd	3rd	4th	5th
Conditional mean/unconditional mean					
U.S. Data	3.01	1.15	0.61	0.30	0.18
Model (HA-T)	3.01	1.05	0.54	0.29	0.11

Note: The source of U.S. data is the Survey of Income and Program Participation 2001.

with the findings in Chang and Kim (2007). The sharp difference in the cross-sectional wealth-employment relationship between Model (HA-T) and Model (HA-N) is due to the presence of government transfers, which mitigates the excessively strong precautionary motive of employment among the poor households in this class of the incomplete markets framework (Yum, 2018).

Lastly, Table 4 shows the joint relationship between income and transfers. Specifically, the reported numbers are the ratio of average progressive-component transfers in each income quintile to the unconditional mean progressive-component transfers. In the U.S., there is a clear negative relationship between the amount of income-security transfers and income. Note that, in the model, this is a complicated equilibrium object, which is shaped not only by the parametric assumption on the nonlinear transfer schedule (5) but also by the endogenous household heterogeneity (which is in turn shaped by consumption-saving and labor supply decisions). Despite the relatively simple functional form (5), we can see that the model does an excellent job of replicating the degree of the transfer progressivity in the U.S.

5 Quantitative results

In this section, we report the main business cycle results and illustrate the mechanism underlying the main quantitative results.

5.1 Business cycle statistics

We first compare business cycle statistics of key macroeconomic variables from simulations of the models to those from the data. We filter all the series using the Hodrick-Prescott filter with a

Table 5: Volatility of aggregate variables

		Model			
	U.S. data	(HA-T)	(HA-N)	(RA-T)	(RA-N)
σ_Y	1.50	1.45	1.39	1.98	1.88
σ_C/σ_Y	0.58	0.22	0.27	0.21	0.24
σ_I/σ_Y	2.96	2.65	2.73	2.82	2.80
σ_L/σ_Y	-	0.61	0.57	-	-
σ_H/σ_Y	0.98	0.83	0.55	0.85	0.82
$\sigma_{Y/H}/\sigma_Y$	0.52	0.41	0.49	0.21	0.24

Note: See Table 1 for the description of the model specifications. Each quarterly variable is logged and detrended using the Hodrick-Prescott filter with a smoothing parameter of 1600. Volatility is measured by the percentage standard deviation of each variable. The U.S. statistics are based on aggregate time-series from 1961Q1 to 2016Q4.

smoothing parameter of 1600. The U.S. data statistics are computed using the aggregate data from 1961Q1 to 2016Q4 (see Appendix B for more details). Table 5 summarizes the cyclical volatility of the key aggregate variables: Y is output, C is consumption, I is investment, L is aggregate efficiency unit of labor, H is aggregate hours, and Y/H is average labor productivity. The volatility is measured by the percentage standard deviation, as is standard in the literature. Except for the output volatility, we report the relative volatility, computed as the absolute volatility of each variable divided by that of output.

The most notable finding in Table 5 is that a very high volatility of aggregate hours observed in U.S. data ($\sigma_H/\sigma_Y = 0.98$) is well accounted for by Model (HA-T). This finding is notable for several reasons. First, note that standard real business cycle models are known to have difficulties in generating a large relative volatility of hours without relying on a low curvature of the utility function (or a high Frisch elasticity). In fact, Models (RA-T) and (RA-N) have the stand-in household whose disutility is linear in aggregate hours. When the utility function features zero curvature in the labor supply, we can see that these models indeed generate a substantial relative volatility of hours (0.85 and 0.82, respectively), as shown in Hansen (1985). It is striking that our baseline model, Model (HA-T), delivers a comparably high volatility of hours (0.83).

Chang and Kim (2006) have demonstrated that a large relative volatility of hours obtained through indivisible labor (Rogerson, 1988) in Hansen (1985) may not be robust to incomplete-

Table 6: Cyclicality of aggregate variables

		Model			
	U.S. data	(HA-T)	(HA-N)	(RA-T)	(RA-N)
$Cor(Y,C)$	0.81	0.80	0.84	0.77	0.81
$Cor(Y,I)$	0.90	0.99	0.99	0.99	0.99
$Cor(Y,L)$	-	0.96	0.95	-	-
$Cor(Y,H)$	0.86	0.92	0.96	0.99	0.99
$Cor(Y,Y/H)$	0.30	0.58	0.95	0.77	0.81
$Cor(H,Y/H)$	-0.23	0.21	0.83	0.66	0.70

Note: See Table 1 for the description of the model specifications. Each quarterly variable is logged and detrended using the Hodrick-Prescott filter with a smoothing parameter of 1600. Cyclicality is measured by the correlation of each variable with output. The statistics are based on aggregate time-series from 1961Q1 to 2016Q4.

markets economies with heterogeneous households.[21] We can see this point also in Table 5, showing that Model (HA-N) is considerably less successful in accounting for the large volatility of hours among the four model specifications using indivisible labor. However, our result from Model (HA-T) suggests that once progressive government transfers are incorporated, the heterogeneous-agent incomplete-markets model may perform as well as the Hansen-Rogerson (stand-in household) economy in terms of a large fluctuation of hours over the business cycle.

Having highlighted the most notable difference across the four models, we also note that there are also interesting differences in the volatility of macroeconomic aggregates. For instance, the volatility of consumption and average labor productivity over the business cycle tends to be more consistent with the data in the heterogeneous-agent models than those in the representative-agent models.[22] In both the heterogeneous-agent and representative-agent models, the introduction of transfers tends to reduce the volatility of consumption. This is not surprising, given the nature of government transfers, which effectively provide insurance against aggregate shocks.

We now move on to the focus of this paper: the cyclicality of macroeconomic variables. The first five rows of Table 6 show the correlations of output with other aggregate variables considered in Table 5. The last row shows the correlation between aggregate hours and labor productivity. As is well known in the literature (e.g., King and Rebelo, 1999), most macroeconomic variables such as

[21] The relative volatility of hours in Chang and Kim (2014) and Takahashi (2014) is very similar to the estimate implied by our (HA-N) model.

[22] Note that in the representative-agent models, average labor productivity is proportional to consumption.

consumption, investment, and aggregate hours are highly procyclical in the U.S. Table 6 shows that the strongly positive correlations with output are fairly well replicated in all model specifications regardless of heterogeneity. Therefore, one may conclude that heterogeneity seems irrelevant, at least in regard to explaining highly procyclical macroeconomic variables over the business cycle.

However, we emphasize that there is a key difference in average labor productivity. In the U.S., average labor productivity does not feature strong procyclicality (i.e., $Cor(Y, Y/H) = 0.30$). A related observation is that the correlation between hours and average labor productivity is even weakly negative (-0.23). In contrast, canonical real business cycle models generate highly procyclical average labor productivity, and thus fail to replicate the cyclicality of average labor productivity, as is well known in the literature. High correlations between output and average labor productivity in Models (RA-T) and (RA-N) (0.77 and 0.81, respectively) manifest this weakness as well.[23]

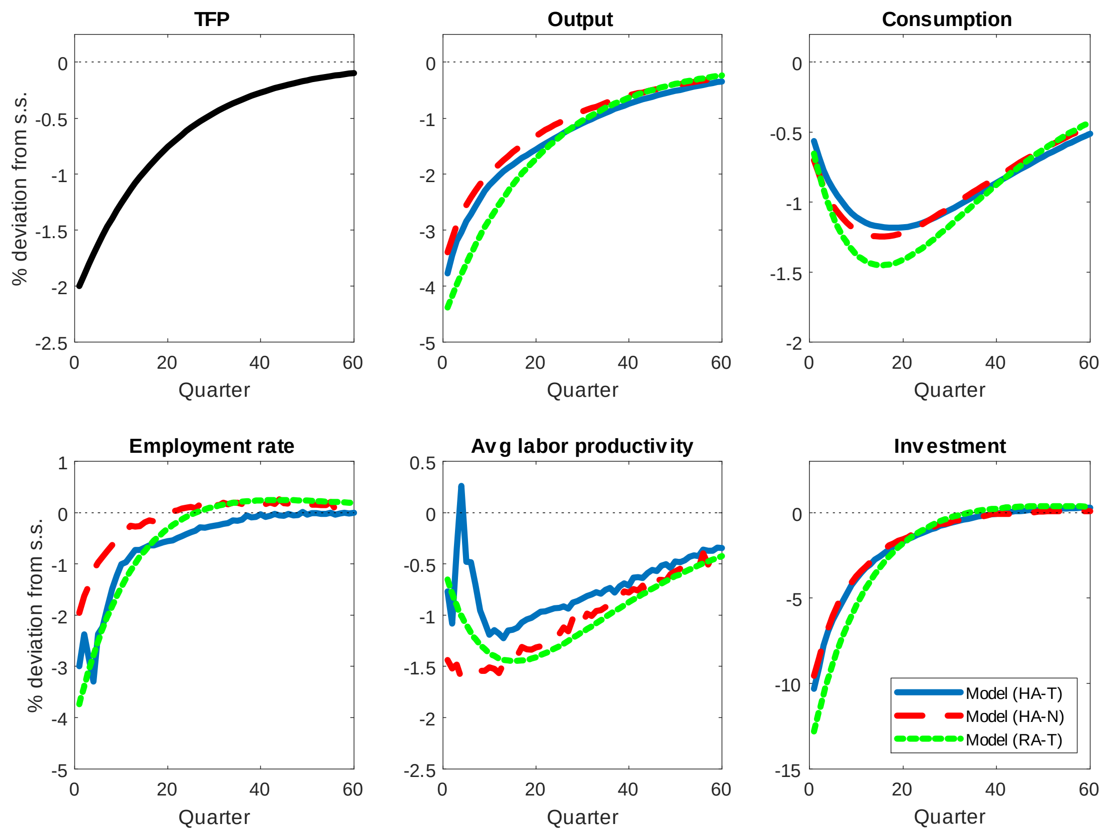
The most notable finding in Table 6 is that the cyclicality of average labor productivity is considerably smaller (0.58) in Model (HA-T), as it is closer to the data (0.30). The literature has suggested various possibilities to dampen strongly positive correlations of average labor productivity with output and hours (Benhabib, Rogerson, and Wright, 1991; Christiano and Eichenbaum, 1992; Braun, 1994; and Takahashi, 2018). In contrast to the existing literature, which relies on additional exogenous shocks, the key to our result is the interaction between household heterogeneity and the presence of government transfers, which generates heterogeneous labor supply behavior across households. In fact, Model (HA-N) which features household heterogeneity, still generates a very high correlation of 0.95, implying that heterogeneity per se cannot dampen highly procyclical average labor productivity in real business cycle models. In the next subsection, we investigate the mechanism underlying our quantitative success.

5.2 Impulse responses

We now present impulse response functions to better understand the main findings in the previous subsection. Figure 3 shows the impulse responses of the key aggregate variables such as output, consumption, aggregate hours, average labor productivity, and investment following a persistent

[23] These correlations would become even higher in models without indivisibility of labor (Hansen, 1985) or without the distortionary labor tax.

Figure 3: Impulse responses of macroeconomic aggregates



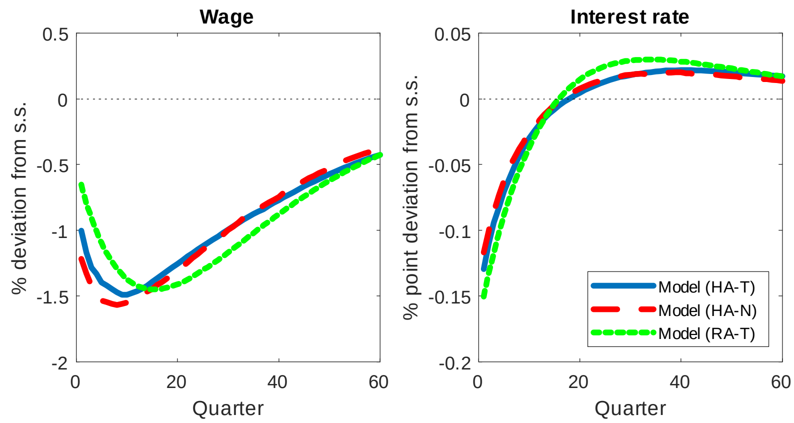
Note: TFP denotes the total factor productivity (or aggregate productivity shocks). The figures display the responses of macroeconomic aggregates to a negative 2 percent TFP shock with persistence ρ_z.

negative 2% shock to z (or TFP) in the first three model specifications.[24] The impulse response of aggregate hours clearly confirms that Model (HA-T) (solid line) delivers a substantially larger fall in hours than Model (HA-N). This reinforces the main finding on a large relative volatility of hours in Model (HA-T) in Table 5. Interestingly, the fall in aggregate hours in Model (HA-T) immediately following a negative TFP shock is not as strong.

Another important difference to note is the impulse responses of average labor productivity. In the representative-agent models, the impulse response of average labor productivity follows

[24] Since the impulse responses from Model (RA-T) and Model (RA-N) are very similar, the figures in this subsection does not report the results from Model (RA-N).

Figure 4: Impulse responses of equilibrium prices



Note: The figures display equilibrium market-clearing price responses to a negative 2 percent TFP shock with persistence ρ_z.

that of consumption, exhibiting an inverse hump shape. In our heterogeneous-agent model without transfers (i.e., Model (HA-N)), the dynamics of average labor productivity follows a more monotone pattern of output, which explains a very high correlation of Y/H with Y in Table 6. In contrast, Model (HA-T) delivers a nontrivial average labor productivity response following a negative TFP shock; it falls initially, and then it reverts to the original level and then it falls again. This nontrivial response of average labor productivity clearly illustrates why the cyclicality of average labor productivity in Model (HA-T) is different from those in the other model specifications.

It is interesting to note that the responses in aggregate output, consumption and investment from Model (HA-T) resemble those from Model (HA-N) very closely and even those from the representative model fairly closely. We now attempt to show how our baseline full model (i.e., Model (HA-T)) delivers the impulse responses of aggregate hours and average labor productivity that are markedly different from the other models. In particular, we focus on illustrating why heterogeneity per se is not sufficient to generate a large response of total hours and non-trivial average labor productivity dynamics.

An obvious candidate is the dynamics of equilibrium prices. Figure 4 displays the changes in market-clearing wage and interest rates following the same negative TFP shock for the first three model specifications. It appears that the difference across the model specifications is not substantial, which suggests that our main results are not driven mainly by the difference in equilibrium price dynamics.

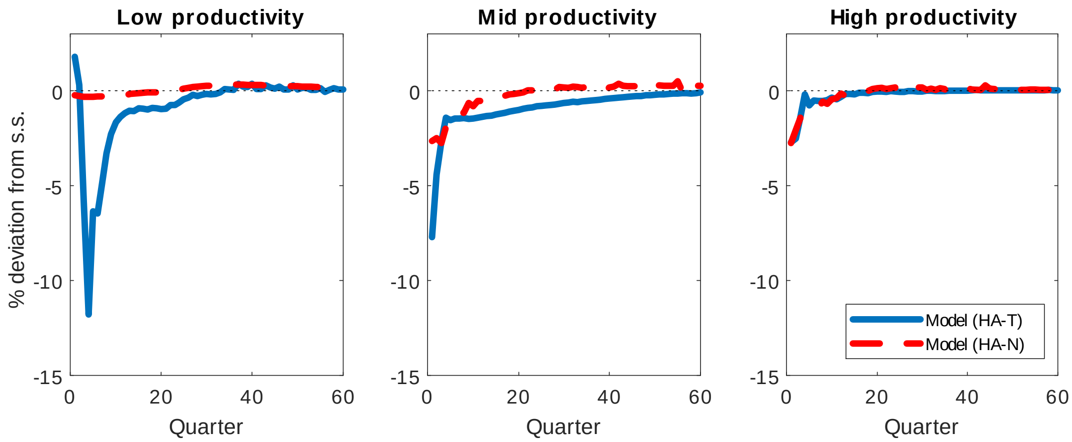
Next, we investigate the impulse responses at a more disaggregated level. Figure 5 plots the impulse responses of hours by productivity. Specifically, we categorize households into three groups: (i) low productivity $\{x_i\}_{i=1}^{4}$; (ii) mid productivity x_5; and (iii) high productivity $\{x_i\}_{i=9}^{6}$. We see that households with higher productivity tend to be less elastic in their labor supply. This result shows that the key insight of Proposition 1 in a simple static model extends to our full heterogeneous-agent model framework. In particular, Model (HA-T) tends to generate a greater difference in the amplitude of hours across productivity, due to the presence of progressive transfers (Proposition 2).

However, we note that there is an exception to this positive relationship between the elasticity and productivity; in Model (HA-N), households with low productivity are *highly inelastic* in their labor supply. The reason for this exceptionally inelastic employment response is related to the finding in Yum (2018), who shows that the absence of public insurance in incomplete markets models raises the precautionary motive for work among wealth-poor households who lack self-insurance. When the precautionary motive is too high, this motive dominates the intertemporal substitution motive, thereby weakening the responses of hours with respect to a persistent fall in wages (see Figure 5). Note also that this inelastic labor supply in the low-productivity group provides a key reason for both a relatively weaker volatility of total hours and a very procyclical average labor productivity in Model (HA-N) relative to Model (HA-T).

6 Microeconomic evidence on heterogeneity in the extensive margin labor supply responses

As shown in the previous sections, the key element of our model is the existence of heterogeneous labor supply responses. More precisely, low-wage workers are considerably more elastic in adjusting labor supply at the extensive margin, which weakens a highly procyclical average labor productivity and, at the same time, enlarges the volatility of aggregate hours worked over the business cycle. In this section, we empirically document heterogeneity in labor supply responses to verify whether our key model mechanism exists in the micro data.

Figure 5: Impulse responses of hours by productivity



Note: Households are grouped into low productivity (below median), mid productivity (median), and high productivity (above median). The figures display employment responses in each group to a negative 2 percent TFP shock with persistence ρ_z.

Specifically, we exploit the panel structure of the PSID to explore whether extensive margin labor supply responses differ by the hourly wage. The panel structure is useful because we can keep track of the same people whose labor supply decisions are observed over time. Because labor supply changes can be measured in different ways and can be shaped by forces at a different level (idiosyncratic vs. aggregate), we consider two approaches. The first approach focuses on identifying the probability of the extensive margin labor supply adjustment for each individual and illustrating how it differs by wage. On the other hand, the second approach focuses on differences in the magnitude of employment rate changes across wage groups during the last six recessions. We present each empirical analysis in more detail.

As mentioned above, the key object of interest in the first approach is the probability of the extensive margin adjustment for each individual. Note that it requires us to have relatively long time-series observations for each individual to obtain a consistent estimate of the adjustment probability, based on the individual-level flow data.[25] Let us fix a year at j. Let i denote an individual index and t denote the year when the individual is observed. We define the extensive margin adjustment based on full-time employment, E, consistent with the previous sections. In other

[25] Since the frequency of the PSID survey has been annual until 1997 and became biannual since 1999, we use the samples observed annually from the 1969-1997 waves.

Table 7: Probability of extensive margin adjustment, by wage quintile

	Length of tracking each individual T		
	5 years	10 years	15 years
Wage quintile in base year			
1st	0.097	0.078	0.068
2nd	0.052	0.044	0.041
3rd	0.039	0.036	0.035
4th	0.035	0.033	0.033
5th	0.040	0.039	0.038
Base years	1969-1993 ($J = 25$)	1969-1988 ($J = 20$)	1969-1983 ($J = 15$)
Avg. no. obs in base years	1,743	1,281	927
Total no. obs.	43,580	25,619	13,911
Avg. age in total sample	43.0	43.7	44.4

Note: See text for the definition of the switching probability reported in this table. Numbers in parentheses show the number of base years.

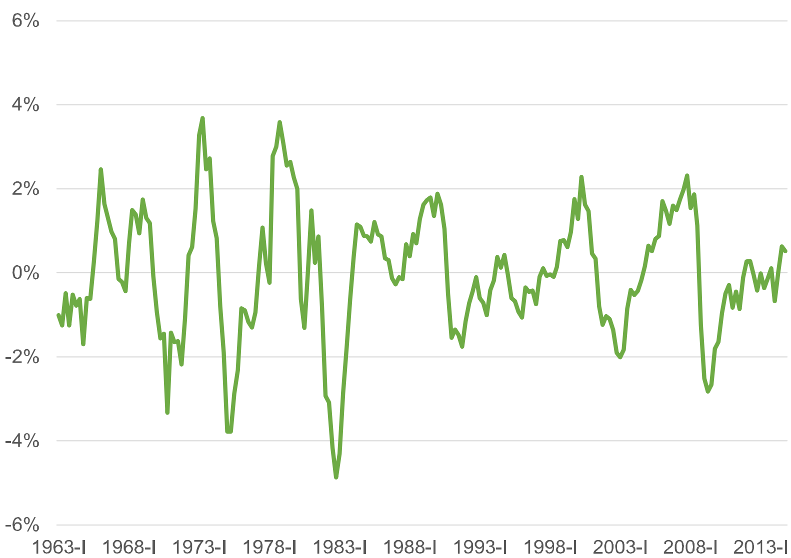
words, an individual i in year t is full-time employed (i.e., $E_{i,t} = 1$) if the annual hours worked are greater than 1,000 hours.[26] Then, we define a binary switching variable, $S_{i,t}$, such that $S_{i,t} = 1$ if $E_{i,t} \neq E_{i,t-1}$ and $S_{i,t} = 0$ otherwise. We exclude the transition from $E_{i,t-1} = 1$ to $E_{i,t} = 0$ if the individual has a non-zero unemployment spell in period t to rule out transitions caused by layoffs.

Note that, given the length of tracking each individual, T, there are $(T-1)$ numbers of $S_{i,t}$ for each individual i. Once we take the average over time, we obtain the individual-specific probability of extensive-margin adjustment at an annual frequency (i.e., $p_{i,j} \equiv \frac{1}{T-1}\sum_{t=j+1}^{j+T-1} S_{i,t}$). As we are interested in differences across wage distribution, we compute p_j^q, defined as the conditional mean of $p_{i,j}$ for each individual's wage quintile bin $q \in \{1, 2, ..., 5\}$ determined in the base year j.

We consider three different values for the length of tracking each individual: $T \in \{5, 10, 15\}$ because a different value of T entails a trade-off. On the one hand, a larger number is beneficial because we are more likely to have a consistent estimate of the adjustment probability at the individual level. On the other hand, a longer time of tracking implies a stricter restriction on samples (because we keep only samples that are observed for T *consecutive* years). Given the value

[26] The results in this section is quite robust to alternative threshold values around $1,000$ for the full time employment.

Figure 6: Cyclical component of real GDP per capita



Note: A quarterly series of real GDP per capita is detrended using HP filter with a smoothing parameter of 1,600.

of T, we compute the estimates of $\{p_j^q\}_{q=1}^5$ by changing the base year j. That way, we attempt to mitigate variations due to the difference in the initial distribution of wage, which is potentially affected by business cycle fluctuations. The reported values in Table 7 are the mean switching probabilities by wage quintile, averaged across the base years, $p^q \equiv \frac{1}{J}\sum p_j^q$ where J is the number of base years (J is reported in parentheses in Table 7).

Table 7 reveals a clear pattern: the individual-level probability of adjusting the labor supply along the extensive margin is significantly higher among low-wage workers. For instance, when $T = 5$, the probability of switching to/from full-time employment among the first wage quintile is 9.7% at the annual frequency. In particular, we can see that this probability tends to decrease with wage. For the third to fifth quintiles, this probability is relatively flat at approximately 4%. When T increases, we also find that the key pattern of extensive margin adjustment probabilities across wage quintiles is still present. Because the samples become slightly older and T becomes longer, however, we also see that the switching probabilities become generally lower.

The above exercise is based on long-run information on the labor market flow at the individual

Table 8: Employment changes in recessions, by wage quintile

	Recession					
	1969-71	1973-76	1980-83	1990-92	2000-02	2006-10
Wage quintile in peak year						
1st	-6.0	-13.9	-11.2	-7.4	-9.8	-17.7
2nd	-6.9	-7.6	-5.3	-8.2	-5.2	-12.9
3rd	-4.6	-7.4	-4.8	-5.3	-3.4	-10.8
4th	-5.4	-4.9	-6.6	-5.4	-4.5	-10.3
5th	-1.7	-5.4	-5.3	-4.4	-1.7	-6.3
No. obs.	1,749	1,838	2,095	2,201	2,970	2,873

Note: The year ranges denote the peak and trough years of each recession. Reported values are percentage point changes in the employment rate by wage quintiles (in the peak year of each recession) using the same set of individuals.

level. The next empirical exercise, on the other hand, uses the differences in magnitude of the employment level changes across wage groups during the recessions. More specifically, we choose six recessions, as evident from Figure 6, which plots the cyclical component of quarterly real GDP per capita.[27] For each recession, we choose a peak year and a trough year, guided by Figure 6. Note that our definition of peak and trough years is limited by the frequency of the PSID because the PSID data set is available annually (until 1997) or biannually (since 1999). Therefore, our choice is also based on the aggregate employment declines in each recession event, according to our micro samples in the PSID. The resulting year combinations for each recession are shown in Table 8.

For each recession, we compute the conditional mean of full-time employment by wage quintile in the peak year: $\frac{1}{N^q_{peak}} \sum_i E^q_{i,peak}$ where N^q_{peak} is the number of observations in the wage quintile bin q in the peak year. Then, we measure the percentage point changes in the employment rate by wage quintile in the corresponding trough year: that is, $\frac{1}{N^q_{peak}} \sum_i \left(E^q_{i,trough} - E^q_{i,peak} \right)$. It is important to note that we keep the set of households in each wage group fixed by assigning a wage quintile to each household in the peak year. That way, our measured changes in employment by wage quintile are not affected by compositional changes, but are rather based on decisions by the same households.

Table 8 clearly shows that the employment rate fell most sharply in the first and second wage

[27] A quarterly series of real GDP per capita is detrended using HP filter with a smoothing parameter of 1,600.

Table 9: Employment changes in recessions excluding samples with unemployment spells, by wage quintile

	Recession				
	1973-76	1980-83	1990-92	2000-02	2006-10
Wage quintile in peak year					
1st	-14.6	-11.5	-7.0	-5.9	-10.8
2nd	-6.6	-2.4	-6.7	-3.6	-7.7
3rd	-7.2	-5.1	-4.4	-2.2	-5.9
4th	-5.8	-7.1	-3.7	-3.7	-6.8
5th	-4.0	-6.7	-4.2	-1.3	-4.6
No. obs.	1,469	1,387	1,689	2,511	2,244

Note: See Table 8 for the basic description. The only difference relative to Table 9 is that we exclude samples that experienced unemployment spells in either the peak year or the trough year. The results for the first recession is omitted because the unemployment information is available only since the 1976 wave (i.e., since the year of 1975).

quintiles in all of the recessions. Furthermore, the magnitude of the decrease in employment tends to be smaller as the wage quintile increases. For example, in the last recession (i.e., the Great Recession), the employment rate among the first wage quintile fell by 17.7 percentage points, whereas the employment rate among the fifth wage quintile fell by only 6.3 percentage points. This pattern of employment changes by wage quintiles is quite robust across different recessions despite variations in the overall changes in the employment rate.[28]

One may be concerned about the possibility that the wage gradient of employment changes found in Table 8 is driven mostly by the firms' demand channel, which may affect household employment status differentially across the wage distribution. To address this concern, we utilize the information about the unemployment spell in the PSID data.[29] More precisely, we exclude samples that experience any unemployment spells over the whole survey years belonging to the range of each recession. That way, we attempt to rule out the effects caused by a differential layoff probability across the wage distribution. Because we impose the additional sample restriction, the number of observations in each recession decreases.

[28] Note that the overall magnitude of the fall in employment is relatively stronger in the recessions of 1973-76, 1980-83 and 2006-10. This finding is, in fact, consistent with relatively larger amplitudes of these recessions, as shown in Figure 6, providing some external validity for our micro samples.

[29] This information is available since the 1976 wave.

Table 9 summarizes the results. In general, we see that the magnitudes of employment changes are somewhat weaker, implying that the demand channel played a role in reducing aggregate employment rates in most recessions. However, we note that our key findings in Table 8 still appears even when we exclude samples that experienced any unemployment spells. First, low-wage workers experience the largest fall in employment across the wage distribution. Furthermore, we still see the general pattern that the magnitude of the fall is negatively related to the wage quintiles in all the recessions. Therefore, we conclude that the salient findings in Table 8 on heterogeneity in employment changes during the recessions remain robust even after accounting for the firms' demand channel.

Although the above two approaches are designed to capture different aspects of labor supply adjustments, they yield consistent results, which demonstrates the robustness of our empirical result that lower-wage workers adjust the labor supply along the extensive margin more elastically. More importantly, both of these empirical findings are consistent with the pattern of heterogeneity in labor supply responses in our model economy, thereby supporting our key mechanism of the heterogeneous-agent model with progressive government transfers.

7 Conclusion

In this paper, we have explored the interplay of household heterogeneity and progressive government transfers in shaping the dynamics of macroeconomic aggregates over the business cycle. We first presented the key insights using analytical results obtained from a stylized static model of the extensive margin labor supply with heterogeneous households. We then constructed a full general equilibrium business cycle model with household heterogeneity. We have shown that in our heterogeneous-agent model with progressive government transfers, which is calibrated to match the micro-level moments, micro-level heterogeneity shapes the dynamics of aggregate labor market variables substantially when household heterogeneity interacts with progressive government transfers. In particular, our baseline real business cycle model delivers moderately positive correlations of average labor productivity with output and hours while generating a large relative volatility of total hours over the business cycle, both of which are difficult to account for by standard real business cycle models.

Using the panel structure of the PSID, we have also documented that the individual-level probability of adjusting the labor supply along the extensive margin is significantly higher among low-wage workers. Furthermore, we have shown that the magnitude of the decline in the employment rate is considerably larger among low-wage workers during the last six recessions. This microeconomic evidence on the heterogeneous responses of labor supply along the extensive margin supports the key mechanism of our heterogeneous-agent model with progressive government transfers.

There are several future research questions that can be naturally followed by our study. First, our analytical result clearly suggests that a more progressive welfare program would reduce the cyclicality of average labor productivity. In the U.S., the correlation between average labor productivity and output has been declining over the last few decades (Galí and van Rens, 2017). At the same time, there has been a steady increase in the size of welfare programs according to the BEA data. It would be very interesting to formally quantify how much of the decrease in the cyclicality of average labor productivity is due to the observed change in the welfare programs. Second, our analysis shows that larger welfare programs would induce a larger fluctuations of aggregate hours, driven mostly by low-wage households. Despite current difficulties in terms of data availability, it would be interesting to test this model implication using cross-country or cross-state data in future work.

References

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf. 2017. "When Inequality Matters for Macro and Macro Matters for Inequality." In *NBER Macroeconomics Annual* 2017, Volume 32: University of Chicago Press.

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *The Quarterly Journal of Economics* 109 (3): 659-684.

Altonji, Joseph G. and Lewis M. Segal. 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business & Economic Statistics* 14 (3): 353-366.

Benhabib, Jess, Richard Rogerson, and Randall Wright. 1991. "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." *Journal of Political Economy* 6: 1166-1187.

Braun, R. Anton. 1994. "Tax Disturbances and Real Economic Activity in the Postwar United States." *Journal of Monetary Economics* 33 (3): 441-462.

Chang, Yongsung and Sun-Bin Kim. 2014. "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations: Reply." *The American Economic Review* 104 (4): 1461-1466.

———. 2007. "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations." *American Economic Review* 97 (5): 1939-1956.

———. 2006. "From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy." *International Economic Review* 47 (1): 1-27.

Chang, Yongsung, Sun-Bin Kim, Kyooho Kwon, and Richard Rogerson. 2019. "Individual and Aggregate Labor Supply in a Heterogeneous Agent Economy with Intensive and Extensive Margins." *International Economic Review,* Forthcoming.

Chang, Yongsung, Sun-Bin Kim, and Frank Schorfheide. 2013. "Labor-Market Heterogeneity, Aggregation, and Policy (in) Variance of DSGE Model Parameters." *Journal of the European Economic Association* 11: 193-220.

Christiano, Lawrence J. and Martin Eichenbaum. 1992. "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations." *American Economic Review* 82 (3): 430-450.

Cociuba, Simona E., Edward C. Prescott, and Alexander Ueberfeldt. 2018. "US Hours at Work." *Economics Letters* 169: 87-90.

Coeurdacier, Nicolas, Hélène Rey, and Pablo Winant. 2015. "Financial integration and growth in a risky world." National Bureau of Economic Research No. w21817.

Den Haan, Wouter J. 2010. "Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents." *Journal of Economic Dynamics and Control* 34 (1): 79-99.

Doepke, Matthias and Michele Tertilt. 2016. "Families in Macroeconomics." In *Handbook of Macroeconomics.* Vol. 2, 1789-1891: Elsevier.

Erosa, Andrés, Luisa Fuster, and Gueorgui Kambourov. 2016. "Towards a Micro-Founded Theory of Aggregate Labour Supply." *The Review of Economic Studies* 83 (3): 1001-1039.

Galí, Jordi and Thijs van Rens. 2017. The Vanishing Procyclicality of Labor Productivity. Working paper.

Hansen, Gary D. 1985. "Indivisible Labor and the Business Cycle." *Journal of Monetary Economics* 16 (3): 309-327.

Heathcote, Jonathan. 2005. "Fiscal Policy with Heterogeneous Agents and Incomplete Markets." *The Review of Economic Studies* 72 (1): 161-188.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2010. "The Macroeconomic Implications of Rising Wage Inequality in the United States." *Journal of Political Economy* 118 (4): 681-722.

———. 2009. "Quantitative Macroeconomics with Heterogeneous Households." *Annual Review of Economics* 1 (1): 319-354.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153-161.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes. 1995. "Precautionary Saving and Social Insurance." *Journal of Political Economy*: 360-399.

Huggett, Mark. 1993. "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies." *Journal of Economic Dynamics and Control* 17 (5): 953-969.

Imrohoroğlu, Ayşe. 1989. "Cost of Business Cycles with Indivisibilities and Liquidity Constraints." *Journal of Political Economy* 97 (6): 1364-1383.

Juhn, Chinhui, Kevin M. Murphy, and Robert H. Topel. 1991. "Why has the Natural Rate of Unemployment Increased Over Time?" *Brookings Papers on Economic Activity* 1991 (2): 75-142.

Khan, Aubhik and Julia K. Thomas. 2008. "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics." *Econometrica* 76 (2): 395-436.

Kim, Heejeong. 2018. "Inequality, Portfolio Choice, and the Business Cycle." Working paper.

King, Robert G. and Sergio T. Rebelo. 1999. "Resuscitating Real Business Cycles." *Handbook of Macroeconomics* 1: 927-1007.

Kopecky, Karen A. and Richard MH Suen. 2010. "Finite State Markov-Chain Approximations to Highly Persistent Processes." *Review of Economic Dynamics* 13 (3): 701-714.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri. 2016. "Macroeconomics and Household Heterogeneity." *Handbook of Macroeconomics* 2: 843-921.

Krusell, Per and Jose-Victor Rios-Rull. 1999. "On the Size of US Government: Political Economy in the Neoclassical Growth Model." *American Economic Review*: 1156-1181.

Krusell, Per and Jr Smith Anthony A. 2006. "Quantitative Macroeconomic Models With Heterogeneous Agents," in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*. Econometric Society Monographs, Vol. 41, ed. by R. Blundell, W. Newey, and T. Persson. Skatteverket: Cambridge University Press, 298-340.

———. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy* 106 (5): 867-896.

Kydland, Finn E. 1984. "Labor-Force Heterogeneity and the Business Cycle." *Carnegie-Rochester Conference Series on Public Policy* 21 (1): 173-208.

Kydland, Finn E. and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica* 50 (6): 1345-1370.

Mendoza, Enrique G., Assaf Razin, and Linda L. Tesar. 1994. "Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption." *Journal of Monetary Economics* 34 (3): 297-323.

Oh, Hyunseung, and Ricardo Reis. "Targeted transfers and the fiscal response to the great recession." *Journal of Monetary Economics* 59 (2012): 50-64.

Rios-Rull, Victor. 1999. "Computation of Equilibria in Heterogeneous Agent Models." In *Computational Methods for the Study of Dynamic Economies*: Oxford University Press.

Rogerson, Richard. 1988. "Indivisible Labor, Lotteries and Equilibrium." *Journal of Monetary Economics* 21 (1): 3-16.

Rouwenhorst, K. Geert. 1995. "Asset pricing implications of equilibrium business cycle models." *In: Cooley, T.F.(Ed.), Frontiers of Business Cycle Research.* Princeton University Press, Princeton, NJ, pp. 294–330.

Takahashi, Shuhei. 2018. "Time-Varying Wage Risk, Incomplete Markets, and Business Cycles." Working paper.

———. 2014. "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations: Comment." *American Economic Review* 104 (4): 1446-1460.

Thomas, Julia K. 2002. "Is Lumpy Investment Relevant for the Business Cycle?" *Journal of Political Economy* 110 (3): 508-534.

Trabandt, Mathias and Harald Uhlig. 2011. "The Laffer Curve Revisited." *Journal of Monetary Economics* 58 (4): 305-327.

Yum, Minchul. 2018. "On the Distribution of Wealth and Employment." *Review of Economic Dynamics* 30: 86-105.

Appendix

A Proofs in Section 2

Proof of Proposition 1 Assume $T_i = 0$. Then, we can rewrite

$$\tilde{a}_i = zx_i.$$

Therefore,

$$N_i = 1 - \exp(-zx_i)$$

Given this, note that

$$\begin{aligned}\varepsilon_i \equiv \frac{\partial N_i}{\partial z}\frac{z}{N_i} &= x_i \exp(-zx_i)\frac{z}{1-\exp(-zx_i)} \\ &= \frac{zx_i \exp(-zx_i)}{1-\exp(-zx_i)}\end{aligned}$$

For expositional convenience, assume that x is continuous for now.

$$\varepsilon(x) = \frac{zx\exp(-zx)}{1-\exp(-zx)}$$

$$\begin{aligned}\frac{\partial\varepsilon(x)}{\partial x} &= \frac{\left[z\exp(-zx) - z^2 x\exp(-zx)\right]\left[1-\exp(-zx)\right] - zx\exp(-zx)\left[z\exp(-zx)\right]}{\left[1-\exp(-zx)\right]^2} \\ &= \frac{\exp(-zx)z\left[1-zx\right]\left[1-\exp(-zx)\right] - z^2 x\exp(-zx)\left[\exp(-zx)\right]}{\left[1-\exp(-zx)\right]^2} \\ &= \frac{z\exp(-zx)\left\{1-zx-\exp(-zx)\right\}}{\left[1-\exp(-zx)\right]^2}\end{aligned}$$

Since $\exp(-zx) < 1$ for all $z, x > 0$,

$$\begin{aligned}\frac{\partial\varepsilon(x)}{\partial x} &= \frac{z\exp(-zx)\left(1-zx-\exp(-zx)\right)}{\left[1-\exp(-zx)\right]^2} < \frac{z\exp(-zx)\left(1-zx-1\right)}{\left[1-\exp(-zx)\right]^2} \\ &= \frac{z\exp(-zx)\left(-zx\right)}{\left[1-\exp(-zx)\right]^2} < 0.\end{aligned}$$

Proof of Proposition 2 Since

$$\frac{\partial N_l}{\partial z} = \exp(-\tilde{a}_l)(1-\lambda),$$
$$\frac{\partial N_h}{\partial z} = \exp(-\tilde{a}_h)(1+\lambda).$$

we have

$$\frac{\partial}{\partial \omega}\left(\frac{\partial N_l}{\partial z}\right) = \exp(-\tilde{a}_l)(1-\lambda)T\lambda > 0,$$
$$\frac{\partial}{\partial \omega}\left(\frac{\partial N_h}{\partial z}\right) = -\exp(-\tilde{a}_h)(1+\lambda)T\lambda < 0.$$

Also, note that

$$\frac{\partial N_l}{\partial \omega} = -\exp(-\tilde{a}_l)T\lambda < 0$$
$$\frac{\partial N_h}{\partial \omega} = \exp(-\tilde{a}_l)T\lambda > 0.$$

Proof of Proposition 3 Since

$$\begin{aligned}\varepsilon &\equiv \frac{\partial N}{\partial z}\frac{z}{N} \\ &= \left(\pi_l \frac{\partial N_l}{\partial z} + \pi_h \frac{\partial N_h}{\partial z}\right)\frac{z}{\pi_l N_l + \pi_h N_h}\end{aligned}$$

the aggregate labor supply elasticity is given by

$$\varepsilon = z\frac{\exp(-\tilde{a}_l)(1-\lambda) + \exp(-\tilde{a}_h)(1+\lambda)}{2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)}$$

where

$$\tilde{a}_l = z(1-\lambda) - T - T\omega\lambda$$

$$\tilde{a}_h = z(1+\lambda) - T + T\omega\lambda.$$

Then, we have

$$\frac{\partial \varepsilon}{\partial \omega} = z \frac{\begin{array}{c}[\exp(-\tilde{a}_l)(1-\lambda)(-1)(-T\lambda) + \exp(-\tilde{a}_h)(1+\lambda)(-1)T\lambda]\,[2-\exp(-\tilde{a}_l)-\exp(-\tilde{a}_h)] \\ -\,[\exp(-\tilde{a}_l)(1-\lambda) + \exp(-\tilde{a}_h)(1+\lambda)]\,[-\exp(-\tilde{a}_l)(-1)(-T\lambda) - \exp(-\tilde{a}_h)(-1)T\lambda]\end{array}}{[2-\exp(-\tilde{a}_l)-\exp(-\tilde{a}_h)]^2}$$

$$= zT\lambda \frac{\begin{array}{c}[\exp(-\tilde{a}_l)(1-\lambda) - \exp(-\tilde{a}_h)(1+\lambda)]\,[2-\exp(-\tilde{a}_l)-\exp(-\tilde{a}_h)] \\ +\,[\exp(-\tilde{a}_l)(1-\lambda) + \exp(-\tilde{a}_h)(1+\lambda)]\,[\exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)]\end{array}}{[2-\exp(-\tilde{a}_l)-\exp(-\tilde{a}_h)]^2}$$

The sign of $\frac{\partial \varepsilon}{\partial \omega}$ is equal to that of the numerator, which can be rewritten as

$$\begin{aligned}
\text{Numerator} = {} & 2(1-\lambda)\exp(-\tilde{a}_l) - (1-\lambda)\exp(-2\tilde{a}_l) - (1-\lambda)\exp(-\tilde{a}_h - \tilde{a}_l) \\
& - 2(1+\lambda)\exp(-\tilde{a}_h) + (1+\lambda)\exp(-\tilde{a}_h - \tilde{a}_l) + (1+\lambda)\exp(-2\tilde{a}_h) \\
& + (1-\lambda)\exp(-2\tilde{a}_l) - (1-\lambda)\exp(-\tilde{a}_h - \tilde{a}_l) \\
& + (1+\lambda)\exp(-\tilde{a}_h - \tilde{a}_l) - (1+\lambda)\exp(-2\tilde{a}_h) \\
= {} & 2\left[(1-\lambda)\exp(-\tilde{a}_l) - (1+\lambda)\exp(-\tilde{a}_h) + 2\lambda\exp(-\tilde{a}_h - \tilde{a}_l)\right].
\end{aligned}$$

Letting $\theta = \frac{(1-\lambda)}{(1+\lambda)}$, we can rewrite

$$\begin{aligned}
& 2(1+\lambda)\left[\frac{(1-\lambda)}{(1+\lambda)}\exp(-\tilde{a}_l) - \exp(-\tilde{a}_h) + \frac{2\lambda}{(1+\lambda)}\exp(-\tilde{a}_h - \tilde{a}_l)\right] \\
& = 2(1+\lambda)\left[\theta\exp(-\tilde{a}_l) + (1-\theta)\exp(-\tilde{a}_h - \tilde{a}_l) - \exp(-\tilde{a}_h)\right].
\end{aligned}$$

Since $\exp(-x)$ is convex, we know

$$\begin{aligned}
\theta\exp(-\tilde{a}_l) + (1-\theta)\exp(-(\tilde{a}_h + \tilde{a}_l)) &> \exp\left(-\left\{\theta\tilde{a}_l + (1-\theta)(\tilde{a}_h + \tilde{a}_l)\right\}\right) \\
&= \exp\left(-\left\{(1-\theta)\,\tilde{a}_h + \tilde{a}_l\right\}\right).
\end{aligned}$$

Applying this inequality, we have

$$\begin{aligned}\text{Numerator} &= 2(1+\lambda)\left[\theta\exp(-\tilde{a}_l) + (1-\theta)\exp(-\tilde{a}_h - \tilde{a}_l) - \exp(-\tilde{a}_h)\right] \\ &> 2(1+\lambda)\left[\exp\left(-\left\{(1-\theta)\,\tilde{a}_h + \tilde{a}_l\right\}\right) - \exp(-\tilde{a}_h)\right] \geq 0\end{aligned}$$

if and only if

$$\begin{aligned}(1-\theta)\,\tilde{a}_h + \tilde{a}_l &\leq \tilde{a}_h \\ \tilde{a}_l &\leq \theta\tilde{a}_h \\ (1+\lambda)\left[z(1-\lambda) - T - T\omega\lambda\right] &\leq (1-\lambda)\left[z(1+\lambda) - T + T\omega\lambda\right] \\ z(1+\lambda)(1-\lambda) - (1+\lambda)T - (1+\lambda)T\omega\lambda &\leq z(1+\lambda)(1-\lambda) - (1-\lambda)T + (1-\lambda)T\omega\lambda \\ -(1+\lambda) - (1+\lambda)\omega\lambda &\leq -(1-\lambda) + (1-\lambda)\omega\lambda \\ -1 &\leq \omega\end{aligned}$$

which is always satisfied.

Proof of Proposition 4 Note that

$$\begin{aligned}\chi_0 &= \frac{(1-\lambda)\,(1-\exp(-\tilde{a}_l)) + (1+\lambda)\,(1-\exp(-\tilde{a}_h))}{2-\exp(-\tilde{a}_l)-\exp(-\tilde{a}_h)} \\ &= \frac{1-\lambda-\exp(-\tilde{a}_l)+\lambda\exp(-\tilde{a}_l)+1+\lambda-\exp(-\tilde{a}_h)-\lambda\exp(-\tilde{a}_h)}{2-\exp(-\tilde{a}_l)-\exp(-\tilde{a}_h)} \\ &= \frac{2-(1-\lambda)\exp(-\tilde{a}_l)-(1+\lambda)\exp(-\tilde{a}_h)}{2-\exp(-\tilde{a}_l)-\exp(-\tilde{a}_h)}.\end{aligned}$$

Therefore, we have

$$
\begin{aligned}
\frac{\partial \chi_0}{\partial z} &= \frac{\left[(1-\lambda)^2 \exp(-\tilde{a}_l) + (1+\lambda)^2 \exp(-\tilde{a}_h)\right]\left[2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right]}{\left(2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right)^2} \\
&- \frac{\left[2 - (1-\lambda)\exp(-\tilde{a}_l) - (1+\lambda)\exp(-\tilde{a}_h)\right]\left[\exp(-\tilde{a}_l)(1-\lambda) + \exp(-\tilde{a}_h)(1+\lambda)\right]}{\left(2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right)^2} \\
&= \frac{1}{\left(2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right)^2}\left\{\begin{array}{c}
2(1-\lambda)^2 \exp(-\tilde{a}_l) + 2(1+\lambda)^2 \exp(-\tilde{a}_h) \\
-(1-\lambda)^2 \exp(-2\tilde{a}_l) - (1+\lambda)^2 \exp(-\tilde{a}_h - \tilde{a}_l) \\
-(1-\lambda)^2 \exp(-\tilde{a}_h - \tilde{a}_l) - (1+\lambda)^2 \exp(-2\tilde{a}_h) \\
-2(1-\lambda)\exp(-\tilde{a}_l) - 2(1+\lambda)\exp(-\tilde{a}_h) \\
+(1-\lambda)^2 \exp(-2\tilde{a}_l) + (1+\lambda)(1-\lambda)\exp(-\tilde{a}_h - \tilde{a}_l) \\
+(1+\lambda)(1-\lambda)\exp(-\tilde{a}_h - \tilde{a}_l) + (1+\lambda)^2 \exp(-2\tilde{a}_h)
\end{array}\right\} \\
&= \frac{2\lambda(\lambda-1)\exp(-\tilde{a}_l) + 2\lambda(\lambda+1)\exp(-\tilde{a}_h) - 4\lambda^2 \exp(-\tilde{a}_h - \tilde{a}_l)}{\left(2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right)^2} \\
&= \frac{2\lambda\left\{(\lambda-1)\exp(-\tilde{a}_l) + (\lambda+1)\exp(-\tilde{a}_h) - 2\lambda\exp(-\tilde{a}_h - \tilde{a}_l)\right\}}{\left(2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right)^2} < 0.
\end{aligned}
$$

Proof of Proposition 5 Define

$$
\Phi(\omega) \equiv \log\left(\frac{\partial \chi_0}{\partial z}\right).
$$

Since the log transformation preserves monotonicity, it suffices to show that $\Phi'(\omega) < 0$. As

$$
\begin{aligned}
\Phi(\omega) &= \log 2\lambda + \log\left\{(\lambda-1)\exp(-\tilde{a}_l) + (\lambda+1)\exp(-\tilde{a}_h) - 2\lambda\exp(-\tilde{a}_h - \tilde{a}_l)\right\} \\
&- 2\log\left(2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right)
\end{aligned}
$$

we have

$$\Phi'(\omega) = \frac{-T\lambda(\lambda-1)\exp(-\tilde{a}_l) + T\lambda(\lambda+1)\exp(-\tilde{a}_h)}{(\lambda-1)\exp(-\tilde{a}_l) + (\lambda+1)\exp(-\tilde{a}_h) - 2\lambda\exp(-\tilde{a}_h - \tilde{a}_l)}$$
$$- 2\frac{T\lambda\exp(-\tilde{a}_l) - T\lambda\exp(-\tilde{a}_h)}{2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)}$$
$$= \frac{\underbrace{T\lambda(1-\lambda)\exp(-\tilde{a}_l) + T\lambda(\lambda+1)\exp(-\tilde{a}_h)}_{\text{positive}}}{\underbrace{(\lambda-1)\exp(-\tilde{a}_l) + (\lambda+1)\exp(-\tilde{a}_h) - 2\lambda\exp(-\tilde{a}_h - \tilde{a}_l)}_{\text{negative}}}$$
$$- 2\frac{\underbrace{T\lambda\left[\exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)\right]}_{\text{positive}}}{\underbrace{2 - \exp(-\tilde{a}_l) - \exp(-\tilde{a}_h)}_{\text{positive}}}$$
$$< 0.$$

B Aggregate data

The business cycle statistics are based on the aggregate time-series data covering from 1961Q1 to 2016Q4. As for output, we use "Real Gross Domestic Product (millions of chained 2012 dollars)" in Table 1.1.6 of the Bureau of Economic Analysis (BEA). As for consumption, we use expenditures in non-durable goods and services reported in Table 2.3.5 of the BEA (Personal Consumption Expenditure). Investment is constructed as the sum of expenditures in durable goods (Table 2.3.5) and private fixed investment in Table 5.3.5. The real values of consumption and investment are calculated using the price index for Gross Domestic Product in Table 1.1.4. Data on total hours worked are obtained from Cociuba et al. (2018). We modified all of the raw time series into those per capita by dividing the raw data by quarterly population in Cociuba et al.(2018).

A target statistic regarding the size of income-security transfers is based on the aggregate data obtained from the BEA. Specifically, we use Supplemental Nutrition Assistance Program (SNAP), Supplemental security income, temporary disability insurance, medical care (Medicaid, general medical assistance and state child health care programs), supplemental security income in Table 3.12 on Government Social Benefits. Note that we do not include large programs such as Medicare, unemployment insurance and veterans' benefits.

C Micro data

For the statistics obtained at the micro level, we use data from the Survey of Income and Program Participation (SIPP). This data set is representative of the non-institutionalized U.S. population. The survey period is in a monthly basis. The SIPP covers a wide range of information on income, wealth, and participation in various transfer programs. We choose the samples from the first wave to the ninth wave of the SIPP in 2001, covering from 2001 to 2003. The original data set is composed of a main module and several topical modules. While the main module contains monthly information on income and transfers, variables such as wealth are reported quarterly in the topical modules. We combine both modules on a quarterly basis.

We construct variables at household level. Data sets in the SIPP contain not only household variables but also individual variables. To generate a household variable from its corresponding individual variable, we take the following steps. First, we identify households with sample unit identifier (SSUID) and household address id in sample unit (SHHADID). Second, we add up the values of a variable for all members in a household. The government transfers that is used to infer the degree of progressivity is based on a broad range of transfer programs including Supplemental Security Income (SSI), Temporary Assistant for Needy Family (TANF), Supplemental Nutrition Assistance Program (SNAP), Supplemental Nutrition Program for Women, Infants, and Children (WIC), childcare subsidy and Medicaid. We do not include age-dependent programs such as Social Security and Medicare. We construct a variable of household income broadly; it consists of labor income, income from financial investments, and property income. We consider households whose head's age is between 23 and 70. We convert all of their nominal values to the values in 2001 US dollar using the CPI-U.

D Estimation of the persistence of full-time worker wage risk

We estimate the persistence of idiosyncratic wage risk in the U.S. using the PSID data, following Heathcote et al. (2010). We choose samples for the period of 1969-2010. Our measure of labor productivity is defined as a worker's relative hourly wage to other individuals. To avoid the oversampling of low income household heads, we exclude households from the Survey of Economic

Opportunity. We consider household heads whose age is between 23 and 65. We drop the samples whose wage is below a half of the minimum wage. The nominal values are converted into the value of US dollar in 2001 with the CPI-U.

We run the ordinary least square regression of the log of the productivity (hourly wages) on a dummy for male, a cubic polynomial in potential experience (age minus years of education minus five), a time dummy, and a time dummy interacted with a college education dummy. We take its residual, $x_{i,j}$, as an idiosyncratic productivity that contains a wide range of individual abilities in the labor market. This stochastic process is composed of the summation of a persistent, $\eta_{i,j}$ and a transitory process, $\nu_{i,j}$:

$$x_{i,j} = \eta_{i,j} + \nu_{i,j}, \nu_{i,j} \sim N(0, \sigma_{\nu}^2) \tag{16}$$
$$\eta_{i,j}^{'} = \rho_{\eta}\eta_{i,j-1} + \epsilon_{i,j}^{'}, \epsilon_{i,j}^{'} \sim N(0, \sigma_{\epsilon}^2)$$

We use a Minimum Distance Estimator to estimate the parameters of the process. The mechanism is to find parameters that minimizing the distance between empirical and theoretical moments. We take the covariance matrix of the residual $x_{i,j}$ as our moments. Let's denote θ as a vector of $(\rho_{\eta}, \sigma_v, \sigma_{\epsilon})$. Let $m_{j,j+n}(\theta)$ be the covariance of the labor productivity between age j and $j+n$ individuals. $\hat{m}_{j,j+n}$ is defined as the empirical counterpart of $m_{j,j+n}(\theta)$. Then,

$$E\left[\hat{m}_{j,j+n} - m_{j,j+n}(\theta)\right] = 0 \tag{17}$$

where

$$\hat{m}_{j,j+n} = \frac{1}{N_{j,j+n}} \sum_{i=1}^{N_{j,j+n}} x_{i,j} \cdot x_{i,j+n}$$

The moments can be represented by as an upper triangle matrix:

$$\bar{m}(\theta) = \begin{bmatrix} m_{0,0}(\theta) & m_{0,1}(\theta) & \cdots & \cdots & m_{0,J-1}(\theta) & m_{0,J}(\theta) \\ 0 & m_{1,1}(\theta) & \cdots & \cdots & m_{1,J-1}(\theta) & m_{1,J}(\theta) \\ 0 & 0 & m_{2,2}(\theta) & \cdots & m_{2,J-1}(\theta) & m_{2,J}(\theta) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & m_{J-1,J-1}(\theta) & m_{J-1,J}(\theta) \\ 0 & 0 & 0 & \cdots & 0 & m_{J,J}(\theta) \end{bmatrix}$$

We denote a vector of $\bar{M}(\theta)$ by vectorizing $\bar{m}(\theta)$ with length $(J+1)(J+2)/2$. To estimate parameters θ, we solve

$$\min_{\theta} \left[\hat{\bar{M}} - \bar{M}(\theta)\right]' W \left[\hat{\bar{M}} - \bar{M}(\theta)\right]$$

where the weighting matrix W is set to be an identity matrix.[30]

E More on calibration for the representative-agent models

It is straightforward to calibrate the parameters of the representative-agent models using the steady state equilibrium equations. First, β is directly obtained by

$$\beta = (1+r)^{-1}$$

Then, given the target of $T/Y = 0.102$ or 0 and $L = 0.750$, B is obtained by

$$B = \frac{(1-\tau)(1-\alpha)}{\left(1 - \delta\frac{K}{Y} - \frac{G}{Y}\right) L}$$

where

$$\frac{K}{Y} = \frac{\alpha}{r+\delta}$$
$$\frac{G}{Y} = \tau(1-\alpha) - \frac{T}{Y}.$$

[30] Using the identity matrix has been common in the literature since Altonji and Segal (1996) show that the optimal weighting matrix generate severe small sample biases.

Finally, since $\frac{Y}{K} = \left(\frac{K}{L}\right)^{\alpha-1}$, we can obtain $\frac{K}{L}$, which in turn gives K and thus Y. Then, T is obtained using $T/Y = .102$.

F More on numerical methods for the heterogeneous-agent models

F.1 Solving for the equilibrium with aggregate risk

The models with aggregate risk are solved in the following two steps. First, we solve for the individual policy functions given the forecasting rules (*the inner loop*). Then, we update the forecasting rules by simulating the economy using the individual policy functions (*the outer loop*). We iterate the two steps until the forecasting rules converge. That is, the difference between the old forecasting rule used in the inner loop and the new forecasting rule generated in the outer loop is small enough.

Inner loop In the inner loop, we solve for the value functions: $V(a, x_i, K, z_k)$, $V^E(a, x_i, K, z_k)$ and $V^N(a, x_i, K, z_k)$, as defined below. These value functions are stored on the non-evenly spaced grid for a and the evenly spaced grid for K with the number of grid points $n_a = 100$ and $n_K = 5$. We use the bivariate cubic spline interpolation along a and K to evaluate the value functions off the grid points. Unlike Chang and Kim (2007; 2014) and Takahashi (2014), we discretize stochastic processes x_i and z_k by using the Rouwenhorst (1995) method. We find that the approximation of continuous AR(1) processes with our estimate of very high persistence is considerably better with the Rouwenhorst method given the same number of grid points.[31] Our baseline results are based on $n_x = 9$ and $n_z = 7$, both of which replicate the true parameters of the continuous AR(1) processes very well. We have solved the model with a greater number of grid points, but the quantitative results change very little. To obtain $V(a, x_i, K, z_k) = \max\left[V^E(a, x_i, K, z_k), V^N(a, x_i, K, z_k)\right]$, we

[31] Specifically, we use the simulated data from Rouwenhorst and Tauchen's methods and estimate the persistence and the standard deviation of error terms in the AR(1) processes for both aggregate productivity shocks and idiosyncratic shocks (available upon request). See also Kopecky and Suen (2010).

solve the following problems

$$V^E(a, x_i, K, z_k) = \max_{\substack{a' > \underline{a}, \\ c > 0}} \left\{ \log c - B\bar{n} + \beta \sum_{j=1}^{N_x} \pi^x_{ij} \sum_{l=1}^{N_z} \pi^z_{kl} V(a', x'_j, \hat{K}', z'_l) \right\} \tag{18}$$

$$\text{subject to} \quad c + a' \leq (1-\tau)\hat{w}(K, z_k)x_i z_k \bar{n} + (1 + \hat{r}(K, z_k))a + T(x_i)$$

and

$$V^N(a, x_i, K, z_k) = \max_{\substack{a' > \underline{a}, \\ c > 0}} \left\{ \log c + \beta \sum_{j=1}^{N_x} \pi^x_{ij} \sum_{l=1}^{N_z} \pi^z_{kl} V(a', x'_j, \hat{K}', z'_l) \right\} \tag{19}$$

$$\text{subject to} \quad c + a' \leq (1 + \hat{r}(K, z_k))a + T(x_i).$$

To evaluate the functional value of the expected value function on $(a', \hat{K}')$ which are not on the grid points, we use the bivariate cubic spline interpolation. By solving these problems, we obtain the individual policy function for work $g_n(a, x_i, K, z_k)$, and that for savings conditional on the labor supply choice:

$$\begin{aligned} g^E_a(a, x_i, K, z_k) = \underset{a' > \underline{a}}{\operatorname{argmax}} \Big\{ & \log\big((1-\tau)\hat{w}(K, z_k)x_i z_k \bar{n} + (1 + \hat{r}(K, z_k))a + T(x_i) - a'\big) \\ & - B\bar{n} + \beta \sum_{j=1}^{N_x} \pi^x_{ij} \sum_{l=1}^{N_z} \pi^z_{kl} V(a', x'_j, \hat{K}', z'_l) \Bigg\} \end{aligned}$$

and

$$\begin{aligned} g^N_a(a, x_i, K, z_k) = \underset{a' > \underline{a}}{\operatorname{argmax}} \Big\{ & \log\big((1 + \hat{r}(K, z_k))a + T(x_i) - a'\big) \\ & + \beta \sum_{j=1}^{N_x} \pi^x_{ij} \sum_{l=1}^{N_z} \pi^z_{kl} V(a', x'_j, \hat{K}', z'_l) \Bigg\}. \end{aligned}$$

Outer loop In the outer loop, we simulate the model economy using the individual policy functions. Two points are worth noting. First, we make sure that $V^E(a, x_i, K, z_k)$ and $V^N(a, x_i, K, z_k)$ satisfy a single-crossing property with respect to a so that there is a unique threshold asset $a^*(x_i, K, z_k)$ for each individual productivity level (conditional on the aggregate state), below which $V^E(a, x_i, K, z_k) > V^N(a, x_i, K, z_k)$ holds and households choose to work. Second, we find the market-clearing prices and the associated aggregate labor in each period during the simulation (Takahashi, 2014).

The measure of households $\mu(a, x_i)$ is approximated by a finer (non-evenly spaced) grid on a than that in the inner loop (Rios-Rull, 1999) with the number of grid points equal to 5000. K is constructed based on the measure of households following $K = \int_a \sum_{i=1}^{N_x} a\mu(da, x_i)$. In each simulation period, we use a bisection method to obtain the equilibrium factor prices as follows:

1. Set an initial range of (w_L, w_H) and calculate the aggregate labor demand $L^d = (1-\alpha)^{\frac{1}{\alpha}}(z_k/w)^{\frac{1}{\alpha}}K$ implied by the firm's FOC for each w. Note that r is obtained by using the relationship $r = z_k^{\frac{1}{\alpha}}\alpha\left(\frac{w}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}} - \delta$, implied jointly by (8) and (9).

2. Calculate the aggregate efficiency unit of labor supply L^s at each w and make sure that the excess labor demand $(L^d - L^s)$ is positive at w_L and it is negative at w_H.

3. Compute $\tilde{w} = \frac{w_L+w_H}{2}$ and obtain $L^d - L^s$ at $\tilde{w}$. If $L^d - L^s > 0$, set $w_L = \tilde{w}$; otherwise, set $w_H = \tilde{w}$.

4. Continue updating (w_L, w_H) until $|w_L - w_H|$ is small enough.

Taking the measure of households $\mu(a, x_i)$, the aggregate state (K, z_k), and factor prices w and r as given, we compute the aggregate efficiency unit of labor supply $L^s(K, z_k)$ by using the threshold asset $a^*(x_i, K, z_k)$ for each individual productivity. Specifically, we solve (18) and (19) given the expected value function in the next period using interpolation. Note that we use the valued function obtained in the inner loop and the forecasting rule (14) for $\hat{K}' = \Gamma(K, z_k)$ which is not on the grid points of K. Then, the individual household decision rules are given by

$$n = g_n(a, x_i, K, z_k) = \begin{cases} \bar{n} & \text{if } a < a^*(x_i, K, z_k), \\ 0 & \text{otherwise,} \end{cases}$$

where $a^*(x_i; K, z)$ is the level such that $V^E(a^*(x_i, K, z), x_i, K, z_k) = V^N(a^*(x_i, K, z), x_i, K, z_k)$. Having $n = g_n(a, x_i, K, z_k)$ for each grid point (a, x_i) on μ at hand, the aggregate efficiency unit of labor supply is obtained by $L^s(K, z_k) = \int_a \sum_{i=1}^{N_x} x_i g_n(a, x_i, K, z_k)\mu(da, x_i)$. After finding the market-clearing prices, we update the measure of households in the next period by using

$$a' = g_a(a, x_i, K, z_k) = \begin{cases} g^E(a, x_i, K, z_k) & \text{if } a < a^*(x_i, K, z_k), \\ g^N(a, x_i, K, z_k) & \text{otherwise,} \end{cases}$$

and the stochastic process for x_i. We simulate the economy for 10,000 periods, as in Khan and Thomas (2008).

Finally, the coefficients $(a_0, a_1, a_2, b_0, b_1, b_2)$ in the forecasting rules

$$\log K' = a_0 + a_1 \log K + a_2 \log z, \tag{20}$$

$$\log w = b_0 + b_1 \log K + a_2 \log z, \tag{21}$$

are updated by ordinary least squares with the simulated sequence of $\{K', w, K, z\}$. Our parametric assumption on the forecasting rules are the same as those in Chang and Kim (2007; 2014) and Takahashi (2014; 2018). We repeat the whole procedure of the inner and outer loops until the coefficients in the forecasting rules converge.

As is clear in the forecasting rules (20) and (21), households predict prices and the future distributions of capital only with the mean capital stock. Therefore, it is important to check whether the equilibrium forecast rules are precise or not. We summarize results for the accuracy of the forecasting rules for the future mean capital stock K' and for the wage w in Table $A1$. It is clear that all R^2 are very high in both specifications of the model. We also present the accuracy statistic proposed by Den Haan (2010). Since our dependent variables are in logs, we multiply the statistics by 100 to interpret them as percentage errors. We find that the mean errors are sufficiently small (considerably less than 0.1 percent) and the maximum errors are also reasonably small ranging around 0.5-0.6 percent for both models.

F.2 Impulse response functions

To compute impulse response functions, we first simulate the economy for a sufficiently long time so that the economy reaches the stochastic steady state (Coeurdacier et al., 2015). Our results use 100 periods for this, and the results do not change with longer periods. Next, we hit the economy

Table A1: Estimates and accuracy of forecasting rules

Model	Dependent variable	Coefficient Const.	$\log K$	$\log z$	R^2	Den-Haan (2010) Mean (%)	Max (%)
Model (HA-T)	$\log K'$	0.1184	0.9533	0.0993	.99998	0.0986	0.5140
	$\log w$	-0.2884	0.4544	0.7014	.99763	0.0983	0.6138
Model (HA-N)	$\log K'$	0.1231	0.9512	0.0983	.99998	0.0894	0.4954
	$\log w$	-0.3643	0.4873	0.7372	.99864	0.0759	0.5153

with an exogenous disturbance to z (-2%), which then follows the AR(1) process with no further shocks. We let the economy evolve according to these shock realizations. The economy is simulated long enough so that it goes back to the original stochastic steady state. In our exercises, 200 periods are long enough for the economy to return to the steady state.

Because we solve the model using a Markov chain for the TFP shocks, it is a non-trivial task to obtain impulse response functions according to its original continuous process. We construct the impulse responses based on the weighted averages using the linear interpolation for z. Specifically, given a value of z, which follows the original AR(1) process, we compute the weight for z by $\omega_z = (z_{i+1} - z)/(z_{i+1} - z_i)$ where $z \in [z_i, z_{i+1}]$ and z_i and z_{i+1} are the two nearest states of the Markov chain. Taking K as given, we calculate the individual decision rules $g_a(a, x_i; K, z_k)$ and $g_n(a, x_i; K, z_k)$ for each $k = i$ and $i + 1$. Note that the market-clearing factor prices are obtained for each k. The individual decision rules and the equilibrium prices are obtained as the weighted averages such as

$$g_a(a, x_i; K, z) = \omega_z g_a(a, x_i; K, z_i) + (1 - \omega_z) g_a(a, x_i; K, z_{i+1}),$$

$$g_n(a, x_i; K, z) = \omega_z g_n(a, x_i; K, z_i) + (1 - \omega_z) g_n(a, x_i; K, z_{i+1}),$$

$$w(K, z) = \omega_z w(K, z_i) + (1 - \omega_z) w(K, z_{i+1}),$$

$$r(K, z) = \omega_z r(K, z_i) + (1 - \omega_z) r(K, z_{i+1}).$$