Heterogeneity, Transfer Progressivity, and Business Cycles*

Youngsoo Jang Takeki Sunakawa Minchul Yum

April 2021

Abstract

This paper studies how transfer progressivity influences aggregate fluctuations when interacting with household heterogeneity. Using a simple static model of the extensive margin labor supply, we analytically characterize how transfer progressivity influences differential labor supply responses to aggregate conditions across heterogeneous households. We then build a quantitative dynamic general equilibrium model with both idiosyncratic and aggregate productivity shocks, and show that it delivers moderately procyclical average labor productivity and a large cyclical volatility of aggregate hours relative to output. Counterfactual exercises show that redistributive policies have very different implications for aggregate fluctuations, depending on whether tax progressivity or transfer progressivity is used. We provide empirical evidence on the heterogeneity of employment responses across the wage distribution, which supports the key mechanism of our model.

Keywords: Progressivity, targeted transfers, extensive margin labor supply, business cycles, redistributive policies

JEL codes: E32, E24, H31, H53, E21

---

*Jang: Shanghai University of Finance and Economics, China, e-mail: jangys724@gmail.com. Sunakawa: Hitotsubashi University, Japan, e-mail: takeki.sunakawa@gmail.com. Yum: University of Mannheim, Germany, e-mail: minchul.yum@uni-mannheim.de. We thank the editor, Kjetil Storesletten, and three anonymous referees for their helpful comments and suggestions that have improved the paper substantially. We are also grateful to Fabian Kindermann, Alex Ludwig, Eunseong Ma and Yikai Wang for their constructive discussions at various stages of this project, and to Árpád Ábrahám, Björn Brügemann, Jay Hong, Pantelis Kazakis, Aubhik Khan, Sun-Bin Kim, Tom Krebs, Dirk Krueger, Per Krusell, Claudio Michelacci, Matthias Meier, Benjamin Moll, Fabrizio Perri, Shuhei Takahashi, Michèle Tertilt, Julia Thomas, and seminar participants at the Bank of Canada, the Bank of Japan, EUI, Federal Reserve Board, Hitotsubashi, KAIST, Korea Univ., Ohio State Univ., Queen Mary Univ. of London, Sogang, UCL, Univ. of Edinburgh, Univ. of Kent, Univ. of Konstanz, Univ. of Mannheim, Univ. of Queensland, Yonsei, and participants at many conferences for comments. Financial support from the German Research Foundation (DFG) through CRC TR 224 (Project A3) and Japan Society for the Promotion of Science through Grant–in–Aid for Scientific Research (KAKENHI) for Young Scientists B (Project 15K20918) is gratefully acknowledged. A previous version of this paper was circulated with the title "Is Household Heterogeneity Important for Business Cycles?".

1 Introduction

In recent years, a large body of literature has investigated the macroeconomic implications of the progressive nature of taxes and transfers that is prevalent in many developed countries.[1] A natural yet relatively unexplored question is how the progressivity of taxes and transfers affects aggregate fluctuations. Given that the sizes of various welfare programs in the United States—e.g., cash transfers, and food, medical, and childcare support to low-income households—have been steadily growing since the 1970s (Ben-Shalom, Moffitt, and Scholz, 2011), it would be particularly timely and relevant to try to improve our understanding of how transfer progressivity influences business cycles.[2] The key question for this paper is how progressive transfers alter the way aggregate shocks are transmitted to the macroeconomy. We ask whether this channel is important for the dynamics of macroeconomic aggregates, in terms not only of volatility (McKay and Reis, 2016), but also their comovement with output.

We begin by considering a simple static model of the extensive margin labor supply, with two types of agents who differ in their potential earnings. These agents have different assets according to a distribution of wealth featuring a high concentration near zero. In this model, we consider higher transfer progressivity as a variation in the transfer schedule, such that agents with low potential earnings (the low type) receive more than those with high potential earnings (the high type), while holding the average amount of transfers constant. We show that higher transfer progressivity induces the labor supply of the low type to respond more strongly to the aggregate shifter. This is because the threshold asset relevant to the employment decision moves closer to zero, around which there is a higher density of (marginal) agents. Consequently, higher transfer progressivity leads to a lower cyclicality of average labor productivity through changes in the composition of workers (Bils, 1985), and potentially to a greater volatility of aggregate hours driven by the low type.[3] Because these analytical results are derived in a highly stylized static environment (necessary to clearly illustrate the mechanism), this simple model is not able to capture any indirect effects of fiscal policy via changes in the distribution of wealth.

To explore the effects of transfer progressivity while allowing for both direct and indirect effects, we then construct a quantitative, dynamic general equilibrium model. We base this on a

[1] The research questions that have been considered include normative ones (such as optimal progressivity) and positive ones (such as the role of progressivity in explaining macroeconomic outcomes). For example, see Conesa, Kitao and Krueger (2009), Heathcote, Storesletten and Violante (2014), Bick and Fuchs-Schündeln (2018), and Guner, Kaygusuz and Ventura (2020) among others.

[2] Interestingly, some recent papers found that progressivity has had no clear trend during the post war period, despite a few drastic ups and downs (e.g., Ferriere and Navarro (2018) who focused on taxes, and Heathcote, Storesletten, and Violante (2020) who considered both taxes and transfers among working-age individuals).

[3] The implications of higher tax progressivity are a bit more involved, and the simple model does not provide a clear-cut prediction regarding the cyclicality of average labor productivity. We provide a detailed discussion on the different forces at work behind the effects of tax progressivity in Appendix B.

standard incomplete markets framework with heterogeneous households who make consumption-savings and extensive margin labor supply decisions in the presence of both idiosyncratic productivity risk and aggregate risk (Chang and Kim, 2006).[4] Our model incorporates progressive taxes and transfers, captured by two separate parsimonious yet flexible nonlinear functions. We calibrate our model economy to match some salient features in the micro-level data, including the degree of progressivity in welfare programs in the Survey of Income and Program Participation (SIPP) data. To better understand the underlying mechanisms, we also consider several nested versions of the baseline model that abstract either from transfers entirely (similar to Chang and Kim, 2007), from differences in transfers across households (similar to Chang, Kim, and Schorfheide, 2013), or from household heterogeneity (similar to Hansen, 1985).

We find that our baseline model delivers aggregate labor market dynamics that differ considerably from the nested ones. First, the baseline model generates considerably lower correlations between average labor productivity and output (at 0.69) compared to all of the nested model—the latter generating much higher values (of between 0.84 and 0.95 as compared to 0.30 in the data). At the same time, the cyclical volatility of aggregate hours relative to output is 0.73 in the baseline model, which is higher than 0.51 and 0.60 in the nested heterogeneous-agent models. This is much closer to the value of 0.80 obtained from its representative-agent counterpart: a version of a Hansen–Rogerson economy that is known to be successful in generating a high volatility of hours.[5]

Transfers play two roles in our baseline model in delivering the above results. The first relates to the theoretical mechanism highlighted in the simple static model. Using impulse response functions at the disaggregated level, we show that low productivity households are more responsive to changes in aggregate shocks in our baseline model, when compared to the model that does not incorporate differences in transfers across households. The second role arises due to risk and market incompleteness. In the absence of any transfers, the labor supply of low productivity households is highly inelastic because they tend to work irrespective of aggregate conditions for precautionary reasons. The presence of transfers mitigates this precautionary motive, thereby raising the responsiveness of their labor supply to aggregate shocks.

We also use our quantitative model to explore how a change in the progressivity of transfers (or of taxes) would affect both the steady state and aggregate fluctuations. We find that increasing

[4] Their model in turn builds on the workhorse incomplete markets general equilibrium model without aggregate risk, as developed by Huggett (1993) and Aiyagari (1994).

[5] Hagedorn and Manovskii (2008) find that in models with search frictions in the labor market, unemployment benefits closer to the potential wage make the value of working closer to the value of being unemployed. This in turn increases labor market volatilities, and we note similarities with our volatility results in this regard. However, we also note that our key model mechanism relies on *heterogeneity* across households in terms of wealth and the size of transfers. This enables us to go beyond labor market volatilities and study issues related to the cyclicality of average labor productivity.

transfer progressivity further reduces the correlation between average labor productivity and output, and raises the volatility of total hours. This is in line with the results from the simple static model. On the other hand, we find that a higher tax progressivity that still induces the same changes, in terms of the observed overall progressivity of taxes and transfers, actually has limited effects on aggregate labor market dynamics. A novel policy implication is that redistributive policies that are meant to be more progressive may have very different business-cycle consequences, depending on the policy tool used to achieve it (i.e., transfers vs. taxes).

Finally, we use micro data from the Panel Study of Income Dynamics (PSID) to empirically explore the heterogeneity of employment changes—a key underlying force at work in our model.[6] We use two different approaches: the first uses individual-level flow data, and the second uses full-time employment rate changes in the short run shaped by aggregate factors.[7] First, we find that the individual-level probability of adjusting labor at the extensive margin is significantly higher among low-wage workers. Second, we document that the full-time employment rate has fallen more in lower wage quintiles during the most recent recessions. Although these two approaches capture employment adjustments over different time horizons, and are shaped by different forcing variables (i.e., idiosyncratic vs. aggregate factors), we find robust evidence that employment in the case of lower-wage workers is more volatile. This provides some support for our key model mechanism.

There has been great interest in incorporating rich micro-level heterogeneity into macroeconomic models in recent decades. Although extensive studies have shown the importance of heterogeneity in accounting for macroeconomic aggregates and equilibrium prices in the absence of aggregate risk (e.g., Huggett, 1993; and Heathcote, 2005), models from the literature that consider aggregate uncertainty have 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; Khan and Thomas, 2008; and Chang and Kim, 2014). Given our main result—that household heterogeneity at the micro level can be important for understanding the dynamics of macroeconomic variables—our paper is broadly in line with recent papers, such as Krueger, Mitman, and Perri (2016), and Ahn, Kaplan, Moll, Winberry, and Wolf (2017). Both of these find that heterogeneity at the micro level is important in shaping the impact of aggregate shocks on macroeconomic variables. Although the distribution of wealth plays an important role in these studies and in our own, they both focus on aggregate consumption via savings decisions,

[6] There is limited empirical evidence of heterogeneity in labor supply responses at the extensive margin across wage groups. See Kydland (1984) and Juhn, Murphy, and Topel (1991) for earlier evidence. Hoynes, Miller, and Schaller (2012) provide evidence of heterogeneity in employment rates across other (potentially related) dimensions, such as race, gender, age, and education.

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

whereas our paper focuses on aggregate labor market dynamics via labor supply choices.

Weak correlations between average labor productivity and output or hours—often referred to as the Dunlop–Tarshis observation—are known to be difficult to explain using standard real business cycle models. The literature has suggested various mechanisms to dampen strongly positive correlations, with earlier studies relying on the introduction of additional shocks to representative-agent models such as home-production technology shocks (Benhabib, Rogerson, and Wright, 1991), government spending shocks (Christiano and Eichenbaum, 1992), and income tax shocks (Braun, 1994). Recently, Takahashi (2020) reduces the correlation between average labor productivity and hours by incorporating uncertainty shocks into into a standard heterogeneous-agent model (Chang and Kim, 2007). Our result is distinct from the existing literature because our mechanism relies on the existence of institutional features leading to heterogeneous responses.

Our quantitative incomplete markets model highlights the effect government transfers have on the precautionary behavior of poor households. An earlier paper by Hubbard et al. (1995) shows that social insurance discourages precautionary savings among low-income households. In a similar manner, using an incomplete-markets model without aggregate uncertainty, Yum (2018) finds that government transfers help bring the employment rate of wealth-poor households closer to the data. This suggests that transfers influence the precautionary employment motive of such households, which in turn is shown to have important implications for the long-run employment effects of labor taxes. Our present results suggest that the existence of progressive transfers in incomplete markets environments also has important implications for the dynamics of macroeconomic aggregates over the business cycle.

The rest of this paper is organized as follows. Section 2 presents the simple static model and present the analytic results on its key mechanism. Section 3 introduces the quantitative dynamic models. Section 4 explains calibration and shows the properties of the quantitative models in stationary equilibrium. Section 5 presents the main quantitative results. Section 6 presents empirical supporting evidence. Section 7 then presents our conclusions.

2 A static model of the extensive margin labor supply

In this section, we present a simple static model of the extensive margin labor supply.[8] The goal of this section is to illustrate the *direct* effects of fiscal policy on aggregate labor market

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

fluctuations in a tractable way. This tractability is achieved by the simplifying assumption that the distribution of wealth is fixed with respect to changes in fiscal policy, and that it is independent of potential earnings. As fiscal policy may change the distribution of wealth, and because this *indirect* effect could potentially affect the theoretical predictions of this section, we will explore more comprehensive effects using a more realistic dynamic model in the subsequent sections.

Our model here considers a continuum of agents in the unit interval. We assume that there are two types of agents with different potential earnings (or wage offers). That is, the individual component of the potential wage 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 the condition $\pi_l + \pi_h = 1$. Agents also differ in their level of asset holdings a_i, and they can choose to either work full-time or not at all: $n_i \in \{0, 1\}$.

The decision problem of each type i is given by:

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

subject to

$$c_i \leq zx_i n_i + a_i + T_i,$$

where c denotes consumption and $b > 0$ captures the disutility of work. We set $b = \log(2) > 0$ without loss of generality. We use z to denote an aggregate shifter of potential earnings, and consider its small perturbations to be the source of aggregate fluctuations. To study the role of transfer progressivity, we allow transfers to depend on the type i (or potential wages)[9].

The above maximization problem characterizes the optimal decision for the discrete employment choice. Specifically, by comparing the utility conditional on working to that on not working, the agent chooses to work if:

$$\log\left(zx_i + T_i + a_i\right) - b \geq \log\left(T_i + a_i\right),$$

or if

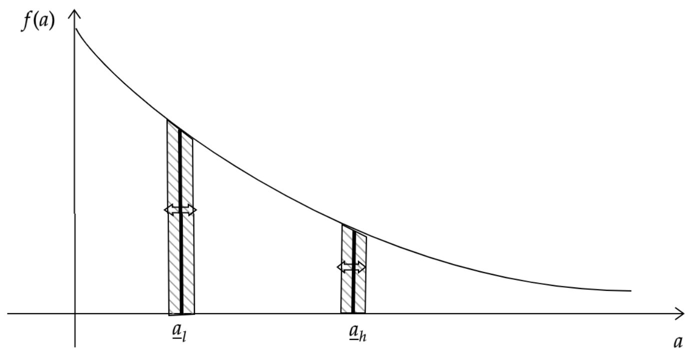
$$a_i \leq zx_i - T_i.$$

This decision rule shows that the agent is more likely to choose to work if the aggregate shifter z or the individual earnings potential x is higher. Also note that the agent is less likely to choose to work if the size of the transfers is higher.

In this model, aggregate employment is determined by both the decision rule and by the

[9] Our focus herein is the difference in transfers between two potential earnings type, and conditional on working (i.e., $T_l > T_h$). The results in this section can also be shown in a generalized environment where transfers when not working is defined separately.

Figure 1: Visual illustration of Proposition 1



Note: The shaded areas denote the relevant density of agents who are affected by perturbations in the aggregate shifter z.

distribution. Let $F(a)$ be the conditional (differentiable) distribution function of assets with its probability density being $f(a) = F'(a)$. Specifically, we use the exponential function in our following results. For $a \geq 0$;

$$F(a) = 1 - \exp(-a),$$
$$f(a) = F'(a) = \exp(-a).$$

This density function has a long right tail in its asset distribution, with a large fraction holding low wealth, being in line with the data.

Given the density function and the decision rule, the fraction of agents working (i.e., the employment rate) for each type is given by:

$$N = F\left(\underline{a}_i\right) = 1 - \exp(-\underline{a}_i),$$

where

$$\underline{a}_i = zx_i - T_i. \tag{1}$$

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

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

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

Assume $T_i = 0$. The labor supply elasticity of agents with low potential earnings is greater than that of agents with high potential earnings. That is, $\varepsilon_l > \varepsilon_h$.

This shows that our model naturally delivers the heterogeneity of labor supply elasticity. The shape of the wealth distribution and the relative location of threshold assets are important for this result. To see this, note that the threshold asset level for the low-potential-wage agents is lower than that for the high-potential-wage agents: $\underline{a}_l < \underline{a}_h$. As shown in Figure 1, the density of the distribution around $\underline{a}_l$ is greater. Since there are more marginal agents around $\underline{a}_l$, the same change in the aggregate shifter z—which perturbs both $\underline{a}_l$ and $\underline{a}_h$—will more strongly affect the employment rate of the low-potential-wage agents.

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.

To study the effects of transfer progressivity, T_i is assumed to be:

$$
\begin{aligned}
T_l &= T\left(1 + \omega\lambda\right), \\
T_h &= T\left(1 - \omega\lambda\right),
\end{aligned}
$$

where $T \in [0, z(1 - \lambda)/2]$ captures the scale of transfers and $\omega \in [0, 1/\lambda]$ shapes the transfer progressivity.[10] Note that a change in ω does not affect the aggregate size of the transfers.[11] Given the above assumptions, the employment rates for each type are given by $N_i = 1 - \exp(-\underline{a}_i)$, where $\underline{a}_l = z(1 - \lambda) - T - T\omega\lambda$ and $\underline{a}_h = z(1 + \lambda) - T + T\omega\lambda$.

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

Intuitively, greater transfer progressivity (or a higher ω) shifts $\underline{a}_l$ to the left where the distribution is denser. There, the same change in the aggregate shifter z would induce more agents to change their employment decision, thereby leading to an even larger elasticity for the low-potential-wage agents. By contrast, greater transfer progressivity shifts $\underline{a}_h$ to the right, around which the distribution of assets is thinner. This implies that the elasticity of the high-potential-wage agents should become smaller.

[10] The maximum values of T and ω ensure that the threshold assets stay non-negative.

[11] As shown in Appendix A, transfers can raise overall progressivity through two channels: (i) the scale of the transfers, and (ii) the relative size of the transfers received by low-income households. A change in ω is meant to capture the second channel, which should better capture welfare transfers.

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 transfer progressivity. That is, $\frac{\partial \varepsilon}{\partial \omega} > 0$.

The key to this result is that an increase in the elasticity of low-potential-wage agents should be large enough to outweigh the opposing effects from a decrease in the elasticity of high-potential-wage types. Given that the density function declines at an increasing rate, this condition is satisfied. This result suggests that transfer progressivity potentially has a role in generating large volatility in aggregate hours, as observed in the data.

Finally, we consider the implications for the cyclicality of average labor productivity. We define average labor productivity as output divided by aggregate hours:

$$\chi \equiv \frac{\sum_{j \in \{l,h\}} \pi_i \left(z x_i N_i \right)}{\sum_{j \in \{l,h\}} \pi_i N_i} = z \frac{\sum_{j \in \{l,h\}} \pi_i \left(x_i N_i \right)}{\sum_{j \in \{l,h\}} \pi_i N_i} \equiv z \chi_0,$$

where we separately define the second term as χ_0. Here, we can clearly see that a change in the aggregate shifter z would directly cause average labor productivity to become procyclical through the first term z, as is the case in real business cycle models. The second term χ_0 captures the effects through worker composition, which indirectly depends on z through heterogeneous employment responses. The following two propositions focus on this second term.

Proposition 4 *A change in the aggregate shifter z has a direct and an indirect effect on average labor productivity $z\chi_0(z)$. The indirect effect is negative: $\frac{\partial \chi_0(z)}{\partial z} < 0$.*

Proposition 5 *Average labor productivity becomes less positively (or more negatively) correlated with z as transfer progressivity increases: $\frac{\partial}{\partial \omega} \left(\frac{\partial \chi_0}{\partial z} \right) < 0$.*

Proposition 5 shows that transfer progressivity can shape the cyclicality of average labor productivity through worker composition effects. To see what this means, let us suppose that the aggregate shifter z increases (i.e., in a boom). While both types of agents are more likely to work, relatively more low-type workers would do so when transfer progressivity is greater. This follows from the disproportionate rise in low-type labor supply elasticity shown in Proposition 3. This force would cause lower increases in average labor productivity during booms, thereby dampening the tight positive link between z and average labor productivity.

Our framework also enables us to explore the implications of tax progressivity. However, the model does not provide a clear prediction when it comes to the cyclicality of average labor productivity. This is because a rise in tax progressivity cause two countervailing effects. On the one hand, the first effect (which moves $\underline{a}_i$) affects the heights of the shaded regions of Figure 1—that is, marginal agents increase. This tends to raise the cyclicality of average labor productivity. On the other hand, the second effect (which changes after-tax earnings or marginal returns to work) instead affects the widths of the shaded regions, and would tend to reduce its cyclicality. We provide a detailed discussion on these two effects in Appendix B.

3 Quantitative business cycle models

As noted earlier, the key results in Section 2 capture the direct effects of fiscal policy changes since they are derived in a static environment. Therefore, it is a quantitative question whether this mechanism would be relevant in a more realistic and dynamic model environment. In the remaining sections, we explore the key mechanisms in models that allow endogenous wealth distributions that can differ by productivity types. We now introduce the economic environment of such quantitative, dynamic general equilibrium models that include aggregate risk.

3.1 Baseline model

The baseline quantitative model we use here builds on a standard incomplete markets framework with both idiosyncratic productivity risk and aggregate risk, as pioneered by Krusell and Smith (1998). In this model, heterogeneous households make a consumption-savings choice—which endogenizes the distribution of wealth—and a labor supply decision at the extensive margin. There are also differences in transfers across households.

Households The model economy is populated by a continuum of infinitely-lived households. It is convenient to describe the decision problem faced by such households in a recursive manner. 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 state i to state j. In addition to the individual state variables, a and x_i, there are also 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 state k to state l. We assume that these Markov processes of individual productivity x and aggregate total factor productivity (TFP) shock z capture the following continuous AR(1) processes in logarithms:

$$\log x' = \rho_x \log x + \varepsilon_x', \tag{2}$$
$$\log z' = \rho_z \log z + \varepsilon_z', \tag{3}$$

where $\varepsilon_x \sim N(0, \sigma_x^2)$ and $\varepsilon_z \sim N(0, \sigma_z^2)$. We denote a variable with a prime symbol 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 also take government policies as given.

The dynamic decision problem facing households can then 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

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

subject to

$$\begin{aligned} c + a' &\leq \tau(e, \bar{e})e + (1 + r(\mu, z_k))a + T \\ e &= w(\mu, z_k)x_i\bar{n} \\ T &= T_1 + T_2(m) \\ m &= e + r(\mu, z_k)\max\{a, 0\} \\ \mu' &= \Gamma(\mu, z_k). \end{aligned} \tag{5}$$

and

$$V^N(a, x_i, \mu, z_k) = \max_{\substack{a' \geq \underline{a}, \\ c \geq 0}} \left\{\log c + \beta \sum_{j=1}^{N_x} \pi_{ij}^x \sum_{l=1}^{N_z} \pi_{kl}^z V(a', x_j', \mu', z_l')\right\} \tag{6}$$

subject to

$$\begin{aligned} c + a' &\leq (1 + r(\mu, z_k))a + T \\ T &= T_1 + T_2(m) \\ m &= r(\mu, z_k)\max\{a, 0\} \\ \mu' &= \Gamma(\mu, z_k). \end{aligned} \tag{7}$$

Households maximize utility by choosing their optimal consumption c, their asset holdings in the next period a', and their labor supply n. Households also face a borrowing limit $\underline{a} \leq 0$. Their labor supply decision is discrete (i.e., $n \in \{0, \bar{n}\}$), and 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 TFP productivity z', as well as the whole distribution μ'. The budget constraints state that the sum of spending should be less than or equal to the sum of income. The evolution of μ is governed by the law of motion, as denoted by $\mu' = \Gamma(\mu, z_k)$.

As shown in the budget constraints, our model incorporates a progressive tax and transfer system, and these two components are captured separately by two nonlinear functions. First, earnings e are subject to progressive taxation—as is standard in the recent quantitative macroeconomics literature. Specifically, for those who have earnings e, progressive taxation leads to a tax rate of:

$$\tau(e, \bar{e}) = \max\left\{1 - \left(\lambda_s \left(e/\bar{e}\right)^{-\lambda_p}\right), 0\right\}. \tag{8}$$

Note that, although this function follows the parametric form of Benabou (2002) and Heathcote, Storesletten and Violante (2014), we restrict $\tau(e)$ to being non-negative. As is well known, $\lambda_p \geq 0$ captures the degree of progressivity and $\lambda_s \geq 0$ inversely controls the scale of taxation. As the input into the progressive tax schedule is earnings normalized by its average $\bar{e}$ (Guner, Kaygusuz and Ventura 2014), a change in λ_p tilts this schedule around average earnings. This strongly affects tax progressivity, yet has little effect on the size of taxation.

On top of this typical progressive tax schedule, we also separately introduce progressive transfers. Following Krusell and Rios-Rull (1999), we make the specific assumption that transfers T consist of two components. 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 the Supplemental Nutrition Assistance Program (SNAP) (formerly known as food stamps), and the Temporary Assistance for Needy Families (formerly the Aid to Families with Dependent Children). As shown in Section 4, the existence of these programs leads us to the observation that the amount of transfers is negatively associated with income. We assume that T_2 depends on total household income m, and use this to replicate the measured transfer progressivity observed in the U.S. data using the following functional form (Yum, 2018):

$$T_2(m) = \omega_s(1 + m)^{-\omega_p}. \tag{9}$$

This parametric assumption adds two parameters. First, $\omega_s \geq 0$ is a scale parameter that determines the overall size of the non-flat part of government transfers (i.e., T_2). The next parameter, $\omega_p \geq 0$, governs the degree of progressivity: a higher ω_p makes T_2 decrease faster with income.

Representative firm and government Aggregate output Y is produced by a representative profit-maximizing firm that solves

$$\max_{K,L} \{z_k F(K,L) - (r(\mu, z_k) + \delta)K - w(\mu, z_k)L\} \tag{10}$$

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 Cobb-Douglas function with constant returns to scale:

$$F(K,L) = K^{\alpha} L^{1-\alpha}. \tag{11}$$

The first-order conditions for K and L give

$$r(\mu, z_k) = z_k F_1(K,L) - \delta, \tag{12}$$

$$w(\mu, z_k) = z_k F_2(K,L). \tag{13}$$

The government in this economy collects labor taxes from households and uses the tax revenue to finance total transfers to households. The remaining tax revenue is spent as government spending G, which is not valued by households. Note that government spending plays no important role in the exercises of this paper.

Equilibrium A recursive competitive equilibrium is a collection of factor prices $r(\mu, z_k)$ and $w(\mu, z_k)$; household decision rules $g_a(a, x_i, \mu, z_k)$ and $g_n(a, x_i, \mu, z_k)$; government spending G; 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 $K(\mu, z_k)$ and $L(\mu, z_k)$; and the aggregate law of motion $\Gamma(\mu, z_k)$; such that

1. Given factor prices $r(\mu, z_k)$ and $w(\mu, z_k)$, the value function $V(a, x_i, \mu, z_k)$ solves the household decision problems defined above, with the associated household decision rules being:

$$a'^{*} = g_a(a, x_i, \mu, z_k), \tag{14}$$

$$n^{*} = g_n(a, x_i, \mu, z_k). \tag{15}$$

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

following (12) and (13).

3. Markets clear:

$$K(\mu, z_k) = \sum_{i=1}^{N_x} \int_a a d\mu \tag{16}$$

$$L(\mu, z_k) = \sum_{i=1}^{N_x} \int_a x_i g_n(a, x_i, \mu, z_k) d\mu. \tag{17}$$

4. Government balances its budget. That is, the sum of government spending G and total transfers to households is equal to the total tax revenue.

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 Alternative model specifications

In addition to the baseline model just introduced, we also consider alternative specifications to illustrate the importance of the interplay between household heterogeneity and government transfers.[12] For convenience, the baseline model featuring "Heterogeneous Agents" and "Targeted" transfers is called Model (HA-T).

The first alternative model specification, denoted as Model (HA-N), is simply a nested specification of the baseline "Heterogeneous-Agent" model with "No" government transfers (i.e., $T_1 = \omega_s = 0$). This model roughly corresponds to the standard incomplete-markets real business cycle model of Chang and Kim (2007), with household heterogeneity and endogenous labor supply at the extensive margin.[13]

The second alternative model specification also keeps household heterogeneity but removes differences in transfers across households. We call this model specification Model (HA-F), which is obtained as a nested "Heterogeneous-Agent" model by making transfers "Flat"—that is, independent of income ($\omega_p = 0$). Chang et al. (2013) also consider a business cycle model that is close to this model specification. Note that this form of transfers (flat lump-sum) is very broadly used in the quantitative macroeconomics literature.

[12] We have also considered a specification which shuts down tax progressivity only. Because its quantitative role is minimal, we have placed those results in Appendix J as a sensitivity check. In Section 5.3, we also consider a counterfactual exercise where we alter tax progressivity using the baseline model specification.

[13] A noticeable difference between Model (HA-N) in our paper and the model in Chang and Kim (2007) is that ours includes progressive taxation whereas theirs does not. However, as shown in Section 5 and Appendix J, the business cycle properties of the model are barely affected by the existence of progressive taxation—except for output volatility.

Our final alternative specification shuts down household heterogeneity. This "Representative-Agent" version of the model is called Model (RA). Given the indivisible labor supply assumption, Model (RA) is essentially the business cycle model studied in Hansen (1985) augmented with taxes and transfers. The key feature of this model specification is that the aggregation of Rogerson (1988) under certain assumptions (such as employment lotteries and consumption insurance) leads to the introduction of the stand-in household whose disutility from work is linear—a powerful mechanism used to generate the large volatility of aggregate hours, observed in U.S. data (Hansen, 1988). Appendix I includes the detailed model environment and its equilibrium definition.

3.3 Solution method

We solve each of the models numerically. Several key features make the numerical solution method nontrivial for the heterogeneous-agent models. First, the 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 a nonlinear method (i.e., the value function iteration) applied to the recursive representation of the problem (as described above). Second, the aggregate law of motion and the state variables involve an infinite-dimensional object: the distribution μ. This requires us to solve the model by approximating the distribution of wealth as its mean (Krusell and Smith, 1998). Since market-clearing is nontrivial in our model with endogenous labor, our solution method also incorporates an additional step to when simulating the model to find market-clearing prices in each period.

We now describe the solution method briefly, with more details found in Appendix I. 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 in order to make current decisions. More precisely, we approximate $\mu(a, x_i)$ by its mean with respect to the asset distribution K. Furthermore, when determining the aggregate capital in the next period K', real wage per efficiency units w and real interest rate r are assumed to be functions of (K, z_k) instead of (μ, z_k). We impose parametric assumptions to approximate the aggregate law of motion $K' = \Gamma(K, z_k)$ and $w = w(K, z_k)$ following

$$\hat{K}' = \hat{\Gamma}(K, z_k) = \exp\left(a_0 + a_1 \log K + a_2 \log z_k\right) \tag{18}$$

$$\hat{w} = \hat{w}(K, z_k) = \exp\left(b_0 + b_1 \log K + b_2 \log z_k\right), \tag{19}$$

as in Chang and Kim (2006, 2007). Households obtain a forecasted $\hat{r}$ based on these forecasting rules, as implied by the first-order conditions of the profit maximization problem facing the representative firm.

The model is solved in two steps. First, given the forecasting rules, we solve for the individual policy functions 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*). It is important to once again note that, since our model environment with endogenous labor supply involves non-trivial factor market clearing, we have to incorporate a step to find the 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 straightforward to solve the representative-agent version of the model. For the purposes of comparison, we keep the same assumptions on the discretization of the TFP shock process as used in the heterogeneous-agent model. The steady-state equilibrium can then be obtained analytically. For solutions with aggregate uncertainty, we use the policy function iteration method.

4 Calibration and model properties in steady state

All model specifications are calibrated to U.S. data. A period in the model is a quarter, as is standard in the business cycle literature. We consider all four of our specifications: Model (HA-T), Model (HA-N), Model (HA-F), and Model (RA).

Calibrating the baseline model We first describe how we calibrate the baseline specification, which involve two sets of parameters. The first set is calibrated externally, in line with the business cycle literature. These parameter values are set in common across all four of our model specifications. The second set of parameters is calibrated to match the same number of relevant target statistics.

We begin by describing the first set of externally calibrated parameters. Most of these are commonly used parameters in the literature. The capital share α is chosen to be consistent with the empirical capital share value of 0.36, and the quarterly depreciate rate δ is set to 2.5%. In our model specifications with a binary labor supply choice, the number of hours worked $\bar{n}$ can be arbitrarily set since it simply determines the scale of the calibrated disutility parameter B. By setting $\bar{n}$ to 1/3—implying that working individuals spend a third of their time endowment on working—we can calibrate $\tilde{B} \equiv B\bar{n}$ directly. Furthermore, the borrowing limit $\underline{a}$ is set to $-T_1/(1+r)$, where r is the equilibrium interest rate in steady state.[14]

[14]This is a form of the natural borrowing limit that ensures that non-working agents are able to pay back their debts in the next period. In Appendix J, we report a version of the model with a zero borrowing limit, as is standard in the literature. The main results found in this paper are robust to this variation.

In the literature, tax progressivity λ_p has been estimated using the same functional form we use. As noted by Holter, Krueger and Stepanchuk (2019), the estimate of λ_p varies quite a lot (from 0.05 to 0.18), depending on the degree of completeness of the data on government transfers used by researchers. Because we model progressive transfers separately in addition to progressive taxes, our taxation parameters in (8) should ideally only capture tax progressivity. As the Internal Revenue Service (IRS) income tax data used by Guner et al. (2014) do not include welfare transfers, we use their estimate for $\lambda_p = 0.053$ and $\lambda_s = 0.911$.[15] As discussed below, we then use micro data on the distribution of welfare transfers across households to calibrate the parameters of the transfer function in (9).

The broad goal of this paper is to study how progressive transfers alter the transmission of aggregate shocks in the macroeconomy. As a first step, we consider the most standard one—total factor productivity shocks (Kydland and Prescott, 1982)—as an aggregate risk, and employ the standard values of $\rho_z = 0.95$ and $\sigma_z = 0.007$ (Cooley and Prescott, 1995). Note that these values are useful as we can easily compare our results to those from recent related papers, such as Chang and Kim (2007) and Takahashi (2020), who also use the same TFP shock estimates.[16]

As the final parameter for this set, ρ_x captures the persistence of idiosyncratic risk in the productivity of households. We estimate the persistence of idiosyncratic risk using the PSID following a standard method from the literature (Heathcote et al., 2010), as discussed in Appendix H. The quarterly value based on this estimate is $\rho_x = 0.9847$. The variability of the idiosyncratic risk is calibrated internally and is explained below. Note that we keep the same values for these two parameters, ρ_x and σ_x, for all of the nested model specifications using heterogeneous agents. This is done in order to control for the underlying idiosyncratic risk present in these models.

The second set of parameters is jointly calibrated. As shown in Table 1, six parameters are calibrated by matching the same number of target statistics. We now explain how each parameter is linked to a target statistic.

The first parameter is $\tilde{B}$, which captures the disutility of work, as defined above. The most relevant target moment is the employment rate of 78.2% from the SIPP sample.[17] The next parameter β captures the discount factor of households. As is standard in the literature, β is

[15] This is the estimate for when the Earned Income Tax Credit (EITC) is included because we do not consider it in our calibration of welfare transfers. We also considered alternative values for λ_p, but these did not affect our quantitative results substantially. This quantitative insignificance of tax progressivity can also be seen explicitly in our counterfactual exercise in Section 5.3.

[16] An interesting exercise for the future would be to investigate how our results might 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 including heterogeneous agents and nonconvexities is an important task, yet is difficult at the present moment due to computational costs.

[17] This value is higher than the employment-population ratio of around 60% that was used in the previous literature (e.g., Chang and Kim, 2007) because we focus here on working-age samples.

Table 1: Parameter values chosen internally

Parameters			Target statistics		
Values		Description	Model	Data	Description
$\tilde{B}=$	.692	Disutility of work	.777	.782	Employment rate
$\beta=$	.985	Subject discount factor	.010	.010	Real interest rate
$\sigma_x=$	.126	S.D. of innovations to $\ln x$	.360	.359	Wage Gini index
$T_1=$	.0337	Overall transfer size	.044	.044	Ratio of Avg (T_1+T_2) to output
$\omega_s=$	.117	Scale of non-flat transfers	.0203	.0201	Ratio of Avg T_2 to output
$\omega_p=$	3.62	Progressivity of transfers	3.07	3.06	$E(T_2 \mid \text{1st income quintile})/E(T_2)$

targeted to match a quarterly interest rate of 1%. The next parameter σ_x governs the variability of idiosyncratic labor productivity. We calibrate this parameter to match the overall wage dispersion captured by the Gini index of worker wages. The target statistic is chosen to be 0.359, which is the average Gini wage in 2000 (Heathcote, Perri and Violante, 2010).[18]

The last three parameters—T_1, ω_s and ω_p—govern the statistics regarding transfers. Recall that T_1 determines the size of universal transfers and ω_s determines the scale of non-flat transfers (T_2). Therefore, the first target statistic regarding transfers is set as the total transfers-output ratio of 4.4%. This is obtained from the time-series average of the ratio of transfers (excluding Social Security and Medicare) to output over the years 1961-2016 according to the Bureau of Economic Analysis (BEA) data. Then, the next target is the average government expenditures on social benefits related to income-security (Table 3.12 from the BEA) over the years 1961–2016—that is, 2.0% of output.[19] Next, we note that ω_p shapes the degree of progressivity in government transfers. Our calibration strategy is to let the model replicate an empirically reasonable degree of transfer progressivity through ω_p, given the value of ω_s. 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 G). Since these welfare programs are highly relevant for poor households, we choose as a target statistic the ratio of the average amount of means-tested transfers received by the first income quintile to its unconditional mean (3.06) (Yum, 2018).

[18] Inequality has been steadily rising in the U.S. In Appendix J, we also consider different values for this target.

[19] We select the components in order to be consistent with our measurement of transfers from the SIPP data, as described below. Our classification of transfers is similar to Krusell and Rios-Rull (1999). See Appendix G for details.

Calibrating alternative model specifications Having explained the calibration strategy of our baseline model, Model (HA-T), we now describe how we calibrate our nested model specifications: Model (HA-N), Model (HA-F), and Model (RA). In general, it would be ideal to minimize the number of parameters to be recalibrated to ensure that our comparison across different model specifications is not driven by different values of parameters. We therefore hold as constant across the nested specifications the parameters governing the idiosyncratic productivity risk: ρ_x and σ_x. However, it is necessary to recalibrate a subset of the internally-calibrated parameters to ensure that the different model specifications are similar in terms of their target statistics (e.g., the employment and interest rates in steady state equilibrium).

Let us first consider Model (HA-N). Because it abstracts from transfers ($T_1 = \omega_s = 0$), the parameter ω_p is irrelevant. We need only to recalibrate $\tilde{B}$ and β to match the employment rate of 78.2% and the real interest rate of 1%. This leads to values of $\tilde{B} = 0.974$ and $\beta = 0.9833$. Next, let us consider Model (HA-F), which shuts down differences in transfers across households. With $\omega_p = 0$, distinguishing between T_1 and ω_s becomes unnecessary. Therefore, we can calibrate the sum of T_1 and ω_s to match the total transfers-output ratio of 4.4%. Aside from this, we recalibrate $\tilde{B}$ and β in the same manner, to yield values of 0.714 and 0.9848, respectively. Finally, and unlike the heterogeneous-agent models, Model (RA) can be calibrated analytically, as is shown in Appendix D. As for the parameters related to tax and transfers, we simply use the average tax rate and the average transfers because progressivity is irrelevant in Model (RA).

Steady state properties Table 1 reports that the baseline model does a good job of matching the target statistics, and the other nested model specifications do a great job of matching a smaller number of targets as well. This does not necessarily mean that the model can account for other relevant statistics. We therefore present the (non-targeted) distributional aspects of the model economy in steady state. First, Table 2 summarizes the share of wealth and the employment rates by wealth quintile from both the model and the data.[20] Overall, all heterogeneous-agent model specifications do a good job of accounting for the shares of wealth held by each wealth quintile.

When we look at the employment rate by wealth quintile reported (also reported in Table 2), we can clearly see 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 (70.0%), and is then relatively flat across the other wealth quintiles. This weak inverted-U-shape of the employment rates across wealth quintiles in the

[20] Table 2 also presents statistics on wealth distribution obtained from the 1992–2007 Survey of Consumer Finances (SCF), as reported by Yum (2018). These statistics from the SCF show a greater concentration of wealth in the top wealth quintile because it better captures the top of the wealth distribution by over-sampling the rich.

Table 2: Characteristics of wealth distribution

	Wealth quintile				
	1st	2nd	3rd	4th	5th
Share of wealth (%)					
U.S. Data (SIPP)	-2.2	1.2	6.8	18.4	76.3
U.S. Data (SCF)	-0.4	1.2	5.1	13.6	80.5
Model (HA-T)	-0.0	0.9	5.2	19.7	74.3
Model (HA-N)	-0.1	0.1	4.8	20.4	74.8
Model (HA-F)	-0.0	0.3	4.9	20.2	74.7
Employment rate (%)					
U.S. Data (SIPP)	70.0	77.9	80.9	82.5	79.7
Model (HA-T)	85.3	79.3	84.4	75.2	64.2
Model (HA-N)	100.0	99.2	74.0	66.0	51.9
Model (HA-F)	100.0	92.0	75.2	67.9	54.0

Note: U.S. data are based on the 2001 Survey of Income and Program Participation (SIPP) and the 1992–2007 Survey of Consumer Finances (SCF) (Yum, 2018). Model (HA-T) is the heterogeneous-agent model with targeted transfers. Model (HA-N) is the heterogeneous-agent model with no transfers. Model (HA-F) is the heterogeneous-agent model with flat transfers.

data is relatively well captured in Model (HA-T). On the other hand, Model (HA-N) predicts that employment falls sharply with wealth, consistent with the findings of Chang and Kim (2007). In this class of the incomplete markets framework, the existence of transfers mitigates the excessively strong precautionary motive for employment among poor households who expect to be near the borrowing limit in the near future (Yum, 2018). Although Model (HA-F) mitigates the negative wealth gradients in employment seen in Model (HA-N), it does not generate the nearly flat employment rates across wealth quintiles.

Next, Table 3 shows the micro relationship between income and transfers in the steady state equilibrium. Specifically, the reported numbers are the ratios of the 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 income and the amount of income-security transfers. Note that, in our model, this is a complicated equilibrium object: it is shaped not only by the parametric assumption on the nonlinear transfer schedule (9) but also by endogenous household choices (such as those regarding consumption-saving and labor supply). Despite the relatively simple functional form (9), we can see that our baseline model does an excellent job of replicating the degree of transfer progressivity in the U.S. Note that, since differences in transfers across households are removed by design in Model (HA-F), this ratio

Table 3: Progressivity of income-security transfers

	Income quintile				
	1st	2nd	3rd	4th	5th
Conditional mean/unconditional mean					
U.S. Data	3.06	0.99	0.52	0.26	0.17
Model (HA-T)	3.07	1.07	0.56	0.24	0.06
Model (HA-F)	1.00	1.00	1.00	1.00	1.00

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

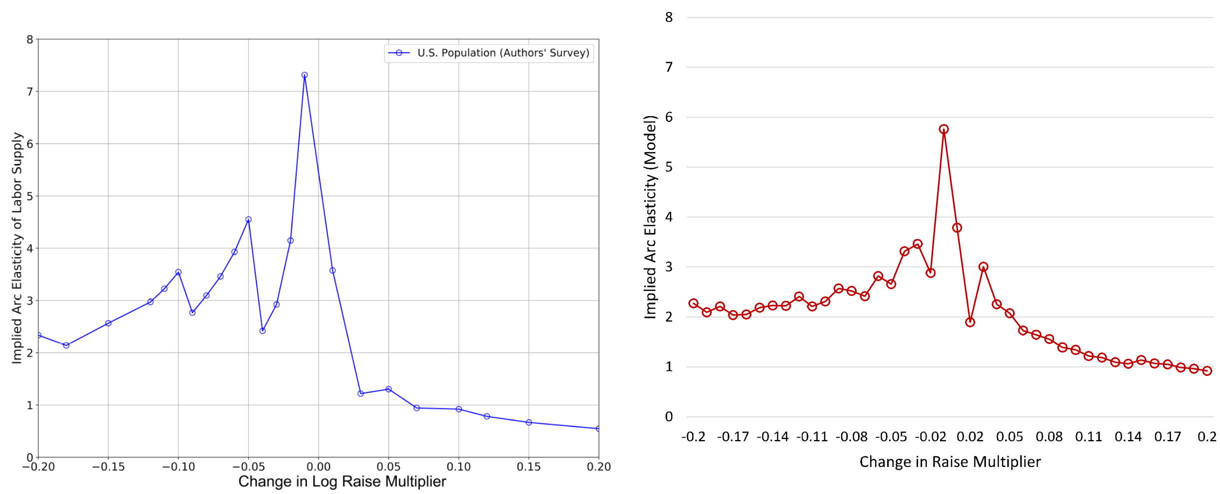
is one for all income quintiles.

We now present our results relating to the aggregate employment responses implied by the model. Given that aggregate employment responses depend on various underlying factors (including the distribution of wealth), they naturally depend on the size of those underlying forcing variables that induce changes in aggregate employment. In this regard, it is very useful to consider the recent empirical evidence of Mui and Schoefer (2020), who propose a novel concept called the reservation raise. In essence, a reservation raise represents a gross percentage change in the (potential) wage of an agent that would make them indifferent between the two choices of working and not working.[21] Once we stack these micro reservation raises, we can obtain an aggregate labor supply curve along the extensive margin, in which the x-axis is the value of reservation raises ranging around one. By constructing arc elasticities around one, the left panel of Figure 2 shows their values as a function of a change in potential earnings using the representative U.S. sample (Mui and Schoefer, 2020), while the right panel plots the model counterpart.

Figure 2 shows that our baseline model replicates two salient patterns observed in the data. As noted by Mui and Schoefer (2020), the empirical arc elasticities show that (i) local elasticities are large with respect to small changes in potential earnings (or reservation raises), and that (ii) elasticities are smaller with respect to large changes in potential earnings. It is worth noting that our baseline model not only qualitatively reproduces these two salient patterns, but it also generates empirically reasonable quantitative responses. Specifically, the elasticities are as high as three (or above) for very small changes in reservation raises (such as $\pm 5\%$). Moreover, they become smaller (at around two) when the change in potential earnings approaches -20%, and they get even smaller again (at around one) when the change in raises approaches +20%. To sum up, it is reassuring to see that our model generates an empirically reasonable aggregate labor

[21] Hence, reservation raises for those who choose to work would be less than one, whereas those for non-workers would be larger than one.

Figure 2: Arc elasticities: data versus model



Note: The left panel is from Mui and Schoefer (2020) who report the U.S. estimates of arc elasticities and the right panel is from the baseline model. The arc elasticities are computed, based on the reservation raise distribution. In essence, a reservation raise represents a gross percentage change in the (potential) wage of an agent that would make them indifferent between the two choices of working and not working.

supply curve, which is a complicated, non-targeted object.[22]

5 Quantitative analysis

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

5.1 Business cycle properties

We first compare the business cycle statistics of key macroeconomic variables from model simulations to those from the data. We filter all the series using the Hodrick–Prescott filter with a smoothing parameter of 1,600. The U.S. data statistics are computed using the aggregate data from 1961Q1 to 2016Q4 (see Appendix F for more details). Table 4 summarizes the cyclical volatility of the following key aggregate variables: output Y, consumption C, investment I, aggregate efficiency unit of labor L, aggregate hours H, and average labor productivity Y/H. Volatility is measured using the percentage standard deviation. As is standard in the business

[22] In Appendix J, we report our results for the other heterogeneous-agent model specifications.

Table 4: Volatility of aggregate variables

		Model			
	U.S. data	(HA-T)	(HA-N)	(HA-F)	(RA)
σ_Y	1.50	1.27	1.48	1.46	1.83
σ_C/σ_Y	0.58	0.27	0.28	0.27	0.25
σ_I/σ_Y	2.96	2.87	2.99	2.99	3.08
σ_L/σ_Y	-	0.50	0.64	0.62	-
σ_H/σ_Y	0.98	0.73	0.51	0.60	0.80
$\sigma_{Y/H}/\sigma_Y$	0.52	0.64	0.54	0.57	0.25

Note: See Table 2 or Section 3.2 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.

cycle literature, our discussion focuses on relative volatility, which is computed as the absolute volatility of each variable divided by that of output.

For several reasons, the most notable finding in Table 4 is that the high volatility of aggregate hours relative to output observed in U.S. data ($\sigma_H/\sigma_Y = 0.98$) is well accounted for by Model (HA-T). We 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). Recall that the disutility of the stand-in household is linear with respect to aggregate hours in Model (RA). When the utility function features zero curvature in the labor supply, we can see that this model indeed generates a substantial relative volatility of hours (0.80), as was also shown by Hansen (1985). It is striking that our baseline model, Model (HA-T), delivers a comparably high volatility of hours (0.73).

In addition, it is important to note from the results of Chang and Kim (2006, 2007) that a large relative volatility of hours obtained through indivisible labor (Rogerson, 1988) in Hansen (1985) may not be robust in incomplete markets economies with heterogeneous households. We can also see this point when we look at the performance of our Model (HA-N), which delivers a substantially smaller volatility of hours (0.51). However, our result from Model (HA-T) suggests that, once heterogeneity in transfers is incorporated in line with the observed patterns in the micro data, the heterogeneous-agent incomplete markets model can perform similarly to the Hansen–Rogerson economy—at least in terms of it having a large relative volatility of hours over the business cycle.

The performance of Model (HA-F) reveals that introducing flat transfers into the model can help obtain a larger relative volatility of hours (0.60). However, this value is still quite far

Table 5: Cyclicality of aggregate variables

		Model			
	U.S. data	(HA-T)	(HA-N)	(HA-F)	(RA)
$Cor(Y,C)$	0.81	0.85	0.85	0.84	0.84
$Cor(Y,I)$	0.90	0.99	0.99	0.99	0.99
$Cor(Y,L)$	-	0.92	0.96	0.96	-
$Cor(Y,H)$	0.86	0.77	0.95	0.87	0.99
$Cor(Y,Y/H)$	0.30	0.69	0.95	0.85	0.84
$Cor(H,Y/H)$	-0.23	0.07	0.81	0.48	0.74

Note: See Table 2 or Section 3.2 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.

from its counterpart in Model (HA-T) of 0.73, suggesting that transfer progressivity achieved by the differences in transfers across households plays an important role. This finding is in line with our analytical results presented in Proposition 3 in Section 2, which shows that greater transfer progressivity increases the degree to which aggregate hours fluctuates with respect to an aggregate shifter (TFP in this case).

Having highlighted the most notable differences across our four model specifications, we would also like to note some of the interesting differences in the volatility of macroeconomic aggregates. For instance, the volatility of average labor productivity over the business cycle tends to be more consistent with the data in the heterogeneous-agent models, as compared to Model (RA). Another observation is that the presence of government transfers tends to reduce the volatility of consumption over the business cycle. This suggests that government transfers play a role as stabilizers, effectively providing insurance against aggregate risk (e.g., see McKay and Reis, 2016).

We now move on to the cyclicality of macroeconomic variables—a key focus of this paper. The first five rows of Table 5 show correlations between output and other aggregate variables. 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 like consumption, investment, and aggregate hours are highly procyclical in the U.S. Table 5 shows that the strongly positive correlations with output are fairly well replicated in all of our model specifications, regardless of the presence of heterogeneity or institutional details. Therefore, one might conclude that heterogeneity or government transfers are irrelevant, at least with respect to the cyclicality of macroeconomic variables over the business cycle.

However, we can see that this conclusion is premature when we look at the comovement of average labor productivity and output. In the U.S., strong procyclicality is not a feature of average labor productivity (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), often referred to as the Dunlop–Tarshis observation (Christiano and Eichenbaum, 1992). By contrast, canonical real business cycle models generate highly procyclical average labor productivity, and thus fail to replicate the limited cyclicality of average labor productivity seen in the data—a problem well-known in the literature. The high correlation between output and average labor productivity in Model (RA) (0.84) is also a manifestation of this weakness.[23]

The most notable finding in Table 5 is that the strong procyclicality of average labor productivity is considerably muted (0.69) in Model (HA-T), and as such it is closer to the data (0.30). In contrast to the existing literature—which tends to rely on the introduction of additional exogenous shocks (e.g., Benhabib, Rogerson, and Wright, 1991; Christiano and Eichenbaum, 1992; Braun, 1994; and Takahashi, 2020)—the key to our result is the interaction between household heterogeneity and transfer progressivity. This in turn generates heterogeneous labor supply behavior across households, as highlighted in Section 2. The importance of the interplay between household heterogeneity and transfers can be seen by the performance of our nested model specifications. Once we abstract from either household heterogeneity or differences in transfers across households, our model generates highly procyclical average labor productivity (above 0.8). In particular, when we abstract from transfers in their entirety (as in Chang and Kim, 2006, 2007) with Model (HA-N), we generate a very high correlation of 0.95. This implies that heterogeneity per se does not dampen highly procyclical average labor productivity in real business cycle models.

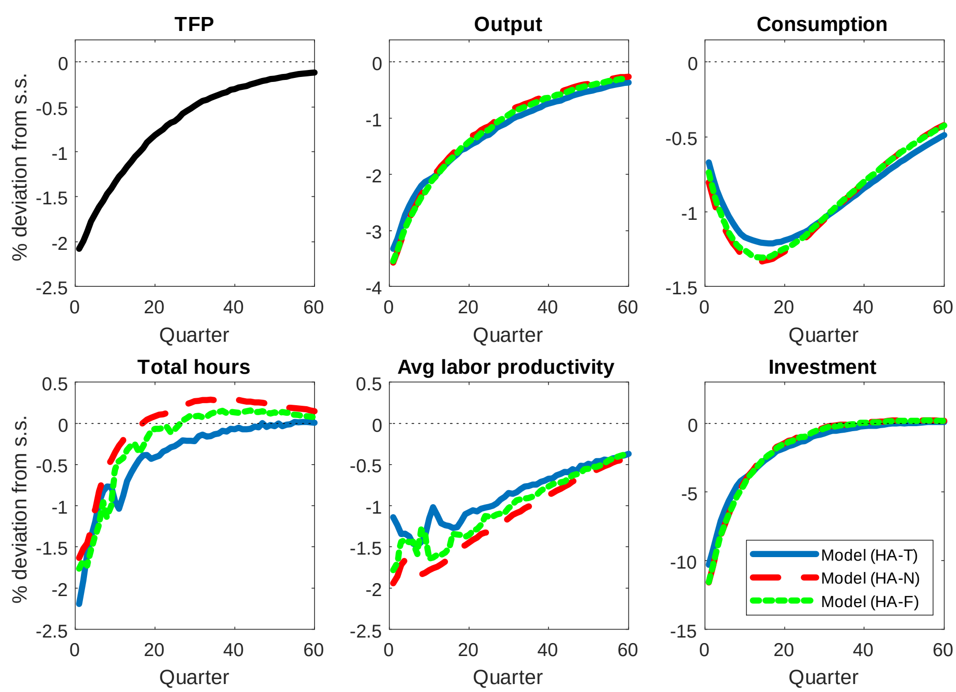
5.2 Impulse responses

We now investigate the mechanism underlying our quantitative success by using impulse response functions. 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 negative 2% shock to z (or TFP) for each of our heterogeneous-agent model specifications. We follow the simulation-based methodology developed by Koop, Pesaran, and Potter (1996), as described in detail in Appendix I (see also Bloom, Floetotto, Jaimovich, Saporta-Eksten and Terry, 2018).

The impulse response of aggregate hours clearly confirms that Model (HA-T) (solid line) de-

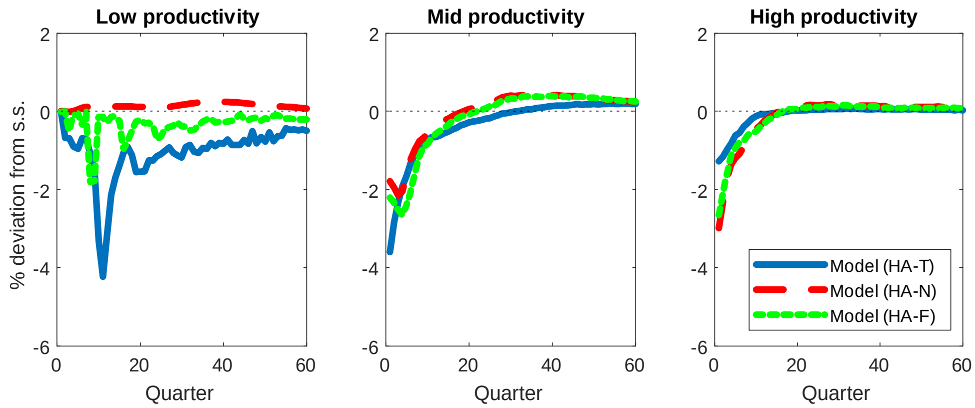
[23] These correlations would become even higher in models without indivisibility of labor (Hansen, 1985) or in the absence of labor taxes.

Figure 3: Impulse responses of macroeconomic aggregates



Note: TFP denotes total factor productivity. The figures display the IRFs of macroeconomic aggregates to a negative 2 percent TFP shock with persistence ρ_z.

Figure 4: Impulse responses of total hours by productivity



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

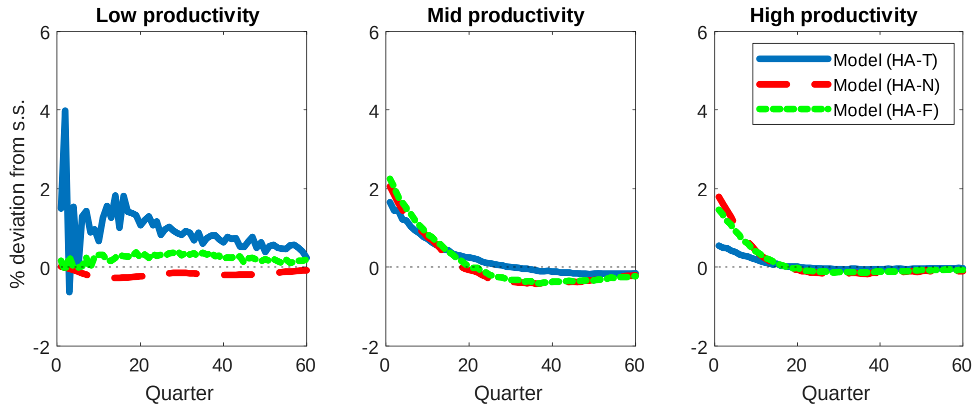
livers a larger fall in hours than the nested heterogeneous-agent models—Model (HA-N) (dashed line) and Model (HA-F) (dotted line)—despite the fact that its output declines the least strongly on impact. These results are consistent with the business cycle results regarding volatility that are presented in Table 4. Another important difference is the impulse responses of average labor productivity. In Model (HA-N), the dynamics of average labor productivity closely follow the pattern of output, since it falls quite sharply on impact. This explains the very high correlation of Y/H with Y in Table 5. When flat transfers are present in Model (HA-F), we see that the overall decrease in average labor productivity is mitigated. In our baseline model, Model (HA-T), the magnitude of the fall in average labor productivity is even smaller, despite it having the largest fall in hours.

To understand the underlying cause of these differences in aggregate dynamics, it is useful to investigate the impulse responses at a more disaggregated level.[24] Specifically, in each period, we categorize households into three almost evenly distributed groups: (i) the low productivity group $\{x_i\}_{i=1}^{4}$; (ii) the mid productivity group $\{x_i\}_{i=5}^{6}$; and (iii) the high productivity group $\{x_i\}_{i=7}^{10}$. Figure 4 plots the impulse responses of total hours by productivity following the same negative shocks, whereas Figure 5 plots its counterparts with respect to positive TFP shocks.

There are several important patterns worth noting. First, there is a relatively small difference

[24]Another obvious candidate is the dynamics of equilibrium prices. Figure A5 displays changes in the market-clearing wage per efficiency units of labor and in real interest rates following the same negative TFP shock for each of our model specifications. It appears as though the difference between these specifications is not substantial among the heterogeneous agent models, suggesting that our main results are not driven mainly by the difference in equilibrium price dynamics.

Figure 5: Impulse responses of total hours by productivity with respect to positive TFP shocks



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

in labor supply responses among the mid productivity group across the different model specifications. On the other hand, the response of the high productivity group is clearly weaker in Model (HA-T) compared to the other heterogeneous-agent models—both the one without differences in transfers across households, Model (HA-F), and the one completely without transfers, Model (HA-N). Second, recall that Proposition 1 from our simple model implies that agents with lower potential earnings tend to be more elastic in their labor supply. In fact, this pattern clearly applies to Model (HA-T), which generates greater magnitudes of changes in labor supply among low productivity groups. This heterogeneity of labor supply responses explains why Model (HA-T) is able to reduce the cyclicality of average labor productivity.

However, this monotonous relationship between elasticity and individual productivity breaks down for the low productivity group, especially in Model (HA-N). This exceptionally inelastic employment response is related to the results of Domeij and Floden (2006) and of Yum (2018). Specifically, both of these papers consider incomplete markets models without public insurance, and show that wealth-poor households who lack self-insurance have precautionary labor supply motives at the intensive margin (Domeij and Floden, 2006) and at the extensive margin (Yum, 2018). Such precautionary motives can dominate the standard intertemporal substitution motive, which in turn could weaken the responses of hours with respect to a persistent fall in wages.

This inelastic labor supply among the low productivity group provides a key reason for both the lower volatility of total hours and the highly procyclical average labor productivity seen in Model (HA-N). This illustrates why heterogeneity per se is not sufficient to explain our key results in incomplete markets environments. The impulse responses from Model (HA-F)—which only shuts down differences in transfers across households—indeed show that the low productivity group has now become responsive to aggregate TFP changes. However, the magnitudes of these responses are still not as sizable as those of Model (HA-T).

Finally, the role of precautionary motives also shows up when we compare Figure 5 to Figure 4. More precisely, Figure 5 shows that, when it comes to positive aggregate shocks, low productivity agents in Model (HA-T) and mid productivity agents in Model (HA-N) and Model (HA-F) do not show muted responses on impact or in several subsequent periods. This is in contrast to their responses following negative aggregate shocks that are present in Figure 4. Given that positive aggregate shocks would move agents away from the borrowing limit and negative shocks toward it, agents are less subject to precautionary motives in these figures.[25]

In summary, the above findings suggest that the presence of progressive transfers in our baseline model plays a dual role. On the one hand, the progressivity of transfers induces low productivity workers to become more responsive to aggregate TFP shocks, as illustrated in the simple model of Section 2. On the other hand, in an incomplete markets environment, the mere presence of transfers helps relax the precautionary motive among wealth-poor households, most of who are low productivity workers. Both roles turn out to be quantitatively important for the dynamics of aggregate hours and average labor productivity.

5.3 Progressivity and the macroeconomy

We now use our baseline model to conduct a counterfactual exercise. Given that there are two separate ways of adjusting the degree of progressivity in the tax-and-transfers system used in our model (i.e., tax progressivity vs. transfer progressivity), we explore the implications of each tool for the macroeconomy. In particular, because our model features aggregate shocks, we investigate their effects not only on steady states but also on aggregate fluctuations. To control for the strength of each policy reform, we make sure that each policy increases the difference between the income Gini coefficients before and after taxes and transfers by 2 percentage points, as compared to the baseline model.[26]

Our results in Table 6 reveal that the steady state effects of higher transfer progressivity differ from those of higher tax progressivity. The former reduces the employment rate of wealth-poor

[25] Figure 5 also reveals that the overall magnitude of employment responses features asymmetry. Specifically, the overall responses are stronger with respect to negative shocks, as compared to those with respect to positive shocks. This is in fact in line with the distribution of arc-elasticities, based on reservation raises: Figure 2 shows that downward adjustments induce greater elasticities (above two) relative to upward adjustments (around one).

[26] Specifically, tax progressivity is increased by raising λ_p by 79% (or $\lambda_p = 0.095$). A higher λ_p tends to raise overall tax revenues, and this works to increase redistribution. As for transfer progressivity, we adjust both ω_p and ω_s simultaneously. This is because a higher ω_p tends to reduce the overall size of transfers, which works against an increase in redistribution. The required percentage increase is 28%.

Table 6: Effects of progressivity on the steady-state economy and aggregate fluctuations

	Baseline Model (HA-T)	Counterfactuals Higher progressivity Transfers	Tax
	Steady state		
Employment rate (%)			
Overall	77.7	71.2	78.4
By wealth quintile			
1st	85.3	50.0	92.7
2nd	79.3	83.8	75.3
3rd	84.4	80.6	85.1
4th	75.2	76.8	74.6
5th	64.2	64.6	64.2
- *Cond. mean/uncond. mean of* T_2 *by income quintile*			
1st	3.07	3.61	3.03
2nd	1.07	0.92	1.09
3rd	0.56	0.34	0.57
4th	0.24	0.12	0.25
5th	0.06	0.02	0.06
	Business cycles		
σ_Y	1.29	1.37	1.29
σ_H/σ_Y	0.73	1.09	0.75
$Cor(Y, Y/H)$	0.69	0.19	0.66
$Cor(H, Y/H)$	0.08	-0.44	0.05

Note: Each counterfactual exercise leads to the same Gini of after-tax-and-transfer income using each policy instrument.

households quite substantially, yet it mildly increases the rate for wealth-rich households. As higher transfer progressivity increases the amount of transfers going to low-income households, the relatively stronger income effects on labor supply can easily explain this result. On the other hand, higher tax progressivity increases the employment rate of the wealth poor. This is because higher progressivity here reduces the tax rate for low-wage households, while raising the tax rate faced by high-wage households.

Having checked the steady state effects, we now move on to the less straightforward implications of these two redistribution policy changes on the business cycle. Table 6 reports that more progressive transfers reduce the cyclicality of average labor productivity and raise the volatility of total hours. These are exactly in line with the main results derived using the simple model in Section 2. On the other hand, higher tax progressivity has limited effects on aggregate labor market dynamics. Average labor productivity becomes slightly less procyclical and the volatility of hours increases marginally. These limited effects of tax progressivity in the quantitative dynamic model are in line with the ambiguous effects of tax progressivity—involving the two countervailing effects—in the simple static model.

Overall, the above results potentially have important policy implications. Policymakers who attempt to pursue more redistributive policies may face very different business cycle consequences, depending on the fiscal tool they use (i.e., taxes vs. transfers).

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

As shown in the previous sections, the key mechanism of our model relies on heterogeneous labor supply responses. More precisely, households with low potential earnings are considerably more elastic in adjusting their labor supply at the extensive margin, which weakens a highly procyclical average labor productivity and 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.

Specifically, we exploit the panel structure of the PSID to explore whether extensive margin labor supply responses differ as a function of hourly wage. This panel structure is useful because we can keep track of the same people and observe their labor supply decisions over time. Because labor supply changes can be measured in different ways and can be shaped by forces at different levels (i.e., 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 illustrates how it differs by wage. Alternatively, the second approach focuses on differences

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

	The length of tracking time T								
	5 years			10 years			15 years		
Wage quintile		Switches			Switches			Switches	
in base year	All	Pos only	Neg only	All	Pos only	Neg only	All	Pos only	Neg only
1st	.097	.061	.036	.075	.048	.027	.066	.042	.024
2nd	.051	.030	.020	.042	.025	.017	.038	.022	.015
3rd	.038	.020	.018	.032	.018	.014	.031	.017	.013
4th	.034	.016	.018	.028	.014	.015	.026	.012	.014
5th	.037	.018	.019	.032	.015	.017	.030	.014	.016
Base years	1969–1993 ($J = 25$)			1969–1988 ($J = 20$)			1969–1983 ($J = 15$)		
Avg. no. obs in base years	1,677			1,189			834		
Total no. obs.	41,920			23,783			12,514		
Avg. age	40.2			41.0			41.5		

Note: See text for the definition of the switching probability reported in this table. Numbers in parentheses show the number of base years. We use samples whose age is between 22 and 64 (inclusive) and who are heads and are not self-employed. "All" refers to the baseline estimates when using both positive and negative switches, whereas "pos only" and "neg only" use only positive ones (i.e., $E_{i,t} = 1$ and $E_{i,t-1} = 0$) and only negative ones (i.e., $E_{i,t} = 0$ and $E_{i,t-1} = 1$), respectively.

in the magnitude of full-time employment rate changes across wage groups during the last six recessions. We now present each empirical analysis in more detail below.

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 this requires us to have relatively long time-series observations for each individual to obtain a consistent estimate of the adjustment probability, based on individual-level flow data.[27] First, let us fix the year at j, and denote i as an individual index, and t the year when the individual is observed. We define the extensive margin adjustment based on full-time employment E: an individual i in year t is in full-time employment (i.e., $E_{i,t} = 1$) if the annual number of hours worked is greater than 1,000.[28] Then, we define a binary switching variable $S_{i,t}$ such that it equals one if $E_{i,t} \neq E_{i,t-1}$ and it equals zero otherwise. We exclude transitions from $E_{i,t-1} = 1$ to $E_{i,t} = 0$ if the individual has a non-zero

[27] Since the frequency of the PSID survey had been annual until 1997 and became biannual from 1999 onward, we use only samples observed annually from the 1969–1997 waves.

[28] The results in this section are quite robust to alternative threshold values for the full-time employment variable. In Appendix K, we report the results when we use 1,500 hours as a full-time threshold value.

unemployment spell in period t in order to rule out transitions caused by layoffs.

Note that, given the length of time spent tracking each individual T, there are $T-1$ counts of $S_{i,t}$ for each individual i. Once we take the average over time, we obtain the individual-specific probability of an extensive-margin adjustment with 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 the wage distribution, we compute p_j^q, which is defined as the conditional mean of $p_{i,j}$ for the wage quintile bin of each individual $q \in \{1, 2, ..., 5\}$ determined in the base year j.

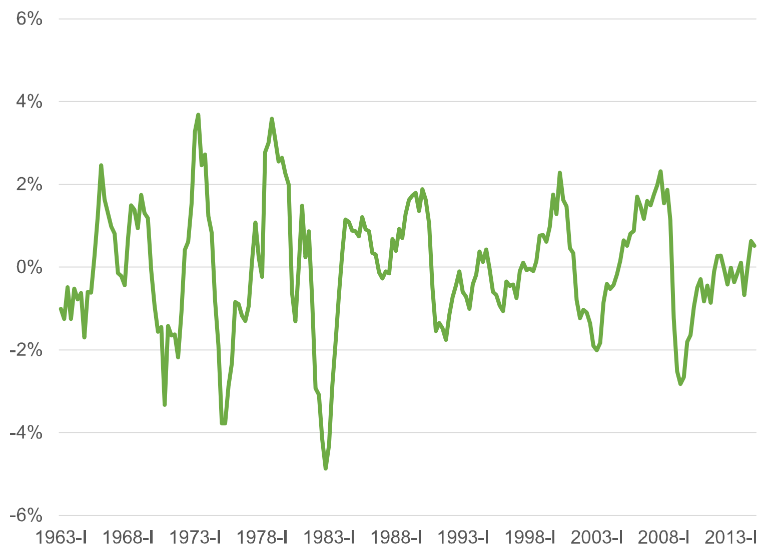
Because different values for the length of time spent tracking each individual entails a trade-off, we consider three variants: $T \in \{5, 10, 15\}$. On the one hand, a larger number is beneficial because we are more likely to have a consistent estimate of the adjustment probability at the individual level. On the other hand, a longer tracking time implies a tighter restriction on the samples—because we keep only samples that were observed for T *consecutive* years. Given the value 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 differences in the initial wage distribution, which is potentially affected by business cycle fluctuations. The reported values in Table 7 are the mean switching probabilities for each wage quintile averaged across the base years, $p^q \equiv \frac{1}{J}\sum p_j^q$, where the number of base years J is reported in parentheses.

Table 7 reveals a clear pattern: the individual-level probability of adjusting the extensive margin is significantly higher among low-wage workers. For instance, when $T = 5$, the probability of switching to or from full-time employment among the first wage quintile is 9.7% (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 3.5%. When T increases, we also find that the key pattern of extensive margin adjustment probabilities across wage quintiles is still present. However, because the samples become slightly older and T becomes longer, we also see that their switching probabilities generally become lower.

We also compute these statistics using only either positive switches ($E_{i,t} - E_{i,t-1} > 0$) or negative switches ($E_{i,t} - E_{i,t-1} < 0$). Table 7 shows that negative wage gradients in the probability of full-time employment adjustments are present in both cases. Interestingly, positive switches feature not only higher adjustment probabilities but also a quantitatively larger negative wage gradient, showing that the overall gradient is more strongly driven by such positive adjustments.

The above exercise is based on long-run information regarding labor market flows at the individual level. The next empirical exercise instead uses the differences in magnitude of full-time employment level changes across wage groups during recessions. More specifically, we choose six recessions and plot the cyclical component of quarterly real GDP per capita in Figure 6. For each recession, we choose a peak year and a trough year, as guided by Figure 6. Note that our definition of peak and trough years is limited by the frequency of the PSID because the data set

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.

was available annually until 1997 and only biannually since 1999. Therefore, our choice is also based on declines in aggregate employment during each recession event—according to our micro samples from the PSID. The resulting year combinations for each recession are shown in Table 8.

Next, we compute the conditional mean of full-time employment by wage quintile in the peak year for each recession by $\frac{1}{N^q_{peak}} \sum_i E^q_{i,peak}$, where N^q_{peak} is the number of observations in wage quintile bin q during the peak year. We then measure the percentage changes in the full-time employment rate by wage quintile in the corresponding trough year. 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 by wage quintile are not affected by compositional changes, but are rather based on changes from within the same households.

Table 8 also clearly shows that the employment rate fell most sharply in the first and second wage quintiles during all of the recessions, and that the magnitude of these declines tends to be smaller among the higher wage quintiles. For example, the full-time employment rate among the first wage quintile during the last recession (i.e., the Great Recession) fell by 17.2%, whereas the counterpart among the fifth wage quintile fell by only 5.8%. This pattern of full-time employment changes by wage quintiles is quite robust across different recessions despite variations in overall

Table 8: Full-time 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	-7.1	-10.1	-9.8	-8.7	-10.0	-17.2
2nd	-3.2	-7.4	-4.5	-6.3	-5.6	-12.9
3rd	-3.7	-7.0	-5.2	-4.6	-3.0	-11.2
4th	-4.7	-4.5	-5.9	-5.4	-4.8	-10.4
5th	-0.9	-5.7	-4.9	-4.6	-2.2	-5.8
No. obs.	1,655	1,756	2,007	2,166	2,924	2,802

Note: The full-time employment threshold is set to 1,000 annual hours. The year ranges denote the peak and trough years of each recession. Reported values are percentage changes in the full-time employment rate by wage quintiles (in the peak year of each recession) following the same set of individuals.

magnitude.[29]

One may be concerned about the possibility that the wage gradient of full-time employment changes found in Table 8 is driven largely by the demand channel of the firms, and that this may affect household employment status differentially across the wage distribution. To alleviate this concern, we now consider the information we have from the PSID data about unemployment spells (available since the 1976 wave or the year of 1975). More precisely, we now exclude samples that experienced any unemployment spells in either the peak year or the trough year. We thereby attempt to rule out the effects caused by differential layoff probabilities across the wage distribution, although the number of observations in each recession decreases because of this additional sample restriction.

Table 9 summarizes our results, whereby we see in general that the magnitudes of full-time employment changes are somewhat weaker. This implies that the demand channel played a sizable role in reducing aggregate employment rates during most recessions. However, we note that our key findings from Table 8 still show up even when we exclude samples that experienced unemployment spells. First, low-wage workers experienced the largest falls in the full-time employment rate across the wage distribution during almost all recessions. Second, we still see the general pattern that the magnitude of this fall is negatively related to wage quintiles during most recessions.

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

Table 9: Full-time 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	-10.7	-4.7	-7.8	-5.2	-8.6
2nd	-5.5	-0.8	-4.9	-3.6	-8.8
3rd	-6.6	-3.4	-4.7	-1.6	-6.4
4th	-4.2	-6.1	-3.5	-4.0	-7.2
5th	-5.2	-5.2	-4.1	-1.8	-4.7
No. obs.	1,547	1,481	1,765	2,454	2,365

Note: The full-time employment threshold is set to 1,000 annual hours. The year ranges denote the peak and trough years of each recession. Reported values are percentage changes in the full-time employment rate by wage quintiles (in the peak year of each recession) following the same set of individuals. Those who experienced unemployment spells in either the peak year or the trough year are excluded. The results for the first recession is omitted because the unemployment information is available only since the 1976 wave (or the year of 1975).

Although the above two approaches are designed to capture different aspects of labor supply adjustments, they yield consistent results: the employment adjustments of lower-wage workers are more elastic. Both of these empirical findings are therefore consistent with the pattern of heterogeneity in labor supply responses seen in the baseline model. Nevertheless, we would also like to stress that the results in this section are only suggestive since we cannot rule out other possible demand-related effects from being behind the observed heterogeneous patterns. For example, they could be consistent with intrasectoral changes in the quality of workers made by firms (Ohanian, 2001). Alternatively, there might be disproportionately more low-wage workers in those industries that are more vulnerable to the effects of the business cycle (Hoynes et al. 2012), or labor hoarding might be differential across wage groups.

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. Using analytical results obtained from a stylized static model of the extensive margin labor supply, we first presented the key insight that higher transfer progressivity would strengthen negative wage gradients in employment responses to aggregate shifters. Using a general equilibrium business cycle model with household heterogeneity, we have shown that micro-level heterogeneity substantially shapes the dynamics of aggregate labor market variables when heterogeneity interacts with progressive transfers. In particular, our baseline model delivers less procyclical average labor productivity when compared to the nested models that are similar to standard real business cycle models. At the same time, it retains the success of the canonical representative-agent indivisible labor supply model in generating a large volatility of aggregate hours without the assumptions of lotteries and perfect consumption insurance (Rogerson, 1988). Our counterfactual analysis showed that greater transfer progressivity lowers the correlation of average labor productivity with output (or hours) even further, and raises the volatility of hours. This is in contrast to greater tax progressivity, which has limited business cycle consequences.

Using the panel structure of the PSID, we also documented how the individual-level probability of experiencing a full-time employment adjustment is significantly higher among low-wage workers. Furthermore, we have shown that the magnitude of the decline in the full-time employment rate was considerably larger among low-wage workers during recent recessions. This microeconomic evidence of the heterogeneous responses of employment reinforces the key mechanism of our heterogeneous-agent model with progressive government transfers.

There are several future research questions that follow naturally from our study. One interesting and novel result that we highlighted in this paper is that the effects of higher transfer progressivity can be quite different from those of higher tax progressivity on both steady state and business cycles—although both can potentially achieve the same observed degree of additional redistribution. A straightforward application relevant to this result would be to design a tax-and-transfer system that also takes into account the welfare costs of business cycles in the presence of heterogeneous agents (e.g., Krusell, Sahin, Mukoyama and Smith, 2009). Although the current paper focused on the transmission of aggregate shocks to the macroeconomy, it would be interesting to explore how other types of aggregate shocks (such as monetary policy shocks) would be transmitted differently in our framework. Finally, our paper introduces some theoretical and quantitative mechanisms suggesting that transfer progressivity might be behind the vanishing procyclicality of average labor productivity, as documented by Galí and van Rens (2021).[30] Formal investigations of these changing relationship are out of the scope of the current paper, but they would nonetheless be highly valuable to address in future work.

[30] There has been a steady increase in the size of welfare programs according to the BEA data. For example, the total government social benefits we considered in our quantitative model (cash transfers, and food, medical, and childcare support to low income households) have increased relative to GDP from 0.5% in the 1960s to 3.5% in the early 2010s.

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.

Benabou, Roland. 2002. "Tax and Education Policy in a heterogeneous-agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70 (2): 481-517.

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

Ben-Shalom, Yonatan, Robert A. Moffitt, and John Karl Scholz. 2011."An Assessment of the Effectiveness of Anti-Poverty Programs in the United States." NBER Working Papers No. 17042.

Bick, Alexander and Nicola Fuchs-Schündeln. 2017. "Taxation and Labour Supply of Married Couples Across Countries: A Macroeconomic Analysis." *The Review of Economic Studies* 85 (3): 1543-1576.

Bils, Mark J. 1985. "Real Wages Over the Business Cycle: Evidence from Panel Data." *Journal of Political Economy* 93 (4): 666-689.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 (3): 1031-1065.

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,* 60 (1): 3-24.

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.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea After all!" *American Economic Review* 99 (1): 25-48.

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

Domeij, David and Martin Floden. 2006. "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased." *Review of Economic Dynamics* 9 (2): 242-262.

Ferrière, Axelle and Gaston Navarro. 2018. "The Heterogeneous Effects of Government Spending: It's all about Taxes." Unpublished Manuscript.

Galí, Jordi and Thijs Van Rens. 2021. "The Vanishing Procyclicality of Labour Productivity." *The Economic Journal* 131 (633): 302-326.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2020. "Child-Related Transfers, Household Labour Supply, and Welfare." *The Review of Economic Studies* 87 (5): 2290-2321.

———. 2014. "Income Taxation of US Households: Facts and Parametric Estimates." *Review of Economic Dynamics* 17 (4): 559-581.

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

Hagedorn, Marcus and Iourii Manovskii. 2008. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited." *American Economic Review* 98 (4): 1692-1706.

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

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

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2020. "How should Tax Progressivity Respond to Rising Income Inequality?" *Journal of the European Economic Association* 18 (6): 2715-2754.

———. 2014. "Optimal Tax Progressivity: An Analytical Framework." *The Quarterly Journal of Economics* 132 (4): 1693-1754.

———. 2010. "The Macroeconomic Implications of Rising Wage Inequality in the United States." *Journal of Political Economy* 118 (4): 681-722.

Holter, Hans A., Dirk Krueger, and Serhiy Stepanchuk. 2019. "How do Tax Progressivity and Household Heterogeneity Affect Laffer Curves?" *Quantitative Economics* 10 (4): 1317-1356.

Hoynes, Hilary, Douglas L. Miller, and Jessamyn Schaller. 2012. "Who Suffers during Recessions?" *The Journal of Economic Perspectives* 26 (3): 27-47.

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.

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.

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

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter. 1996. "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics* 74 (1): 119-147.

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

Krusell, Per, Toshihiko Mukoyama, Ayşegul Şahin, and Anthony A. Smith Jr. 2009. "Revisiting the Welfare Effects of Eliminating Business Cycles." *Review of Economic Dynamics* 12 (3): 393-404.

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. 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.

McKay, Alisdair and Ricardo Reis. 2016. "The Role of Automatic Stabilizers in the US Business Cycle." *Econometrica* 84 (1): 141-194.

Mui, Preston and Benjamin Schoefer. 2020. "Reservation Raises: The Aggregate Labor Supply Curve at the Extensive Margin." Unpublished Manuscript.

Ohanian, Lee E. 2001. "Why did Productivity Fall so Much during the Great Depression?" *American Economic Review* 91 (2): 34-38.

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

Takahashi, Shuhei. 2020. "Time-Varying Wage Risk, Incomplete Markets, and Business Cycles." *Review of Economic Dynamics* 37: 195-213.

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

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

Appendix

A Tax and transfer progressivity

It is customary to define progressivity based on both taxes and transfers (Heathcote et al., 2014). As our paper emphasizes the distinction between tax progressivity and transfer progressivity, we formally define them herein.

First, let us quickly review the standard case based on both taxes and transfers. Conditional on a person working, let $y \geq 0$ denote their income before taxes and transfers (pre-government income), and $G(y)$ denote taxes net of transfers for those whose pre-government income is y. Hence, post-government income is given by $y - G(y)$. It is then easy to show that the average tax rate (ATR) is given by $G(y)/y$, and the marginal tax rate (MTR) is given by $G'(y)$. A tax-and-transfer system is *progressive* if the ATR increases with y, or if $MTR(y) > ATR(y)$.

Now consider the case when taxes and transfers are distinguished explicitly, as in the current paper. Let $\tau(y)$ be the *tax rate* faced by those with pre-government income y such that $\tau(y) \geq 0$ and $T(y)$ be *transfers* to those with pre-government income y such that $T(y) \geq 0$. We can relate this to the above standard case in the following way. First, taxes net of transfers are given by $G(y) = \tau(y)y - T(y)$. Then, since the ATR is:

$$\frac{G(y)}{y} = \frac{\tau(y)y - T(y)}{y} = \tau(y) - \frac{T(y)}{y},$$

and the MTR is:

$$\frac{dG(y)}{dy} = \tau'(y)y + \tau(y) - T'(y),$$

we can show that the overall tax-and-transfer system is progressive if

$$\begin{aligned}
&MTR(y) - ATR(y) \\
&= \tau'(y)y + \tau(y) - T'(y) - \tau(y) + \frac{T(y)}{y} \\
&= \underbrace{\tau'(y)y}_{\text{Tax}} + \underbrace{\frac{T(y)}{y}}_{\text{Transfer size}} \underbrace{-T'(y)}_{\text{Transfer slope}} > 0.
\end{aligned} \tag{A1}$$

In (A1), we can clearly see that overall progressivity in the tax-and-transfer system can be decomposed into one aspect of *tax progressivity* and two aspects of *transfer progressivity*. The first condition controls tax progressivity, as governed by $\tau'(y)$, and the next two terms concern progressivity due to transfers. Specifically, the second term captures the effects coming from the scale of the transfers. As its value is non-negative, this effect always contributes to overall progressivity—and increasingly more so as transfers become more sizable. The final term captures the importance of the slope of the transfer schedule, which is the focus of this paper, and this term contributes to overall progressivity if $T'(y) < 0$—and increasingly more so as $T'(y)$ becomes more negative.

For example, a popular environment of a flat tax and lump-sum (flat) transfers is *progressive*, because the second term is progressive (even though the other two terms are neutral) according to (A1).

B Effects of tax progressivity in the simple model

Our framework allows us to explore the direct effects of tax progressivity. Given that we have two potential wages, and labor supply is only at the extensive margin (no intensive margin adjustment), progressive taxation is captured by $\tau_l < \tau_h$ where after-tax earnings are given by $(1-\tau_i)zx_in_i$. Note that we consider *small* perturbations of z so that its impact on τ_i is of second-order importance (as compared to the difference in τ_i itself). In this stylized model environment, tax progressivity increases when τ_h increases relative to τ_l. One can then easily show that the asset thresholds for employment for each type i are given by:

$$\underline{a}_i = (1-\tau_i)zx_i - T_i.$$

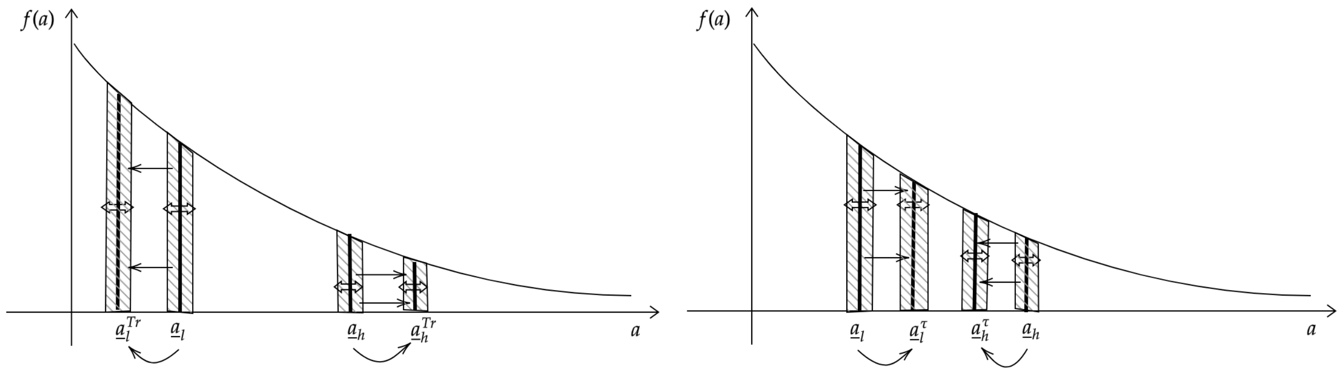
Note that higher tax progressivity moves $\underline{a}_h$ to the left, whereas higher transfer progressivity moves $\underline{a}_h$ to the right. This appears to suggest that the two ways of increasing progressivity (either by decreasing T_h/T_l or by increasing τ_h/τ_l) would have exactly opposite effects on these employment thresholds, as illustrated in the right panel of Figure A1.

However, it is important to note that a change in tax progressivity also affects the width of the shaded regions—that is, the density of marginal agents relevant to the extensive margin elasticity. To see this, it is useful to consider a stark example (i.e., a toy model version of our exercise in Section 5.3). Consider a simple model economy which features both progressive transfers ($T_l \equiv T > 0 = T_h$) and progressive taxes ($\tau_h \equiv \tau > 0 = \tau_l$). Asset thresholds for each i are then given by:

$$\underline{a}_l = x_l - T,$$
$$\underline{a}_h = (1-\tau)x_h.$$

Now we imagine two reforms, as visualized in Figure A2, both of which would increase the

Figure A1: Direct effects of transfer progressivity or tax progressivity



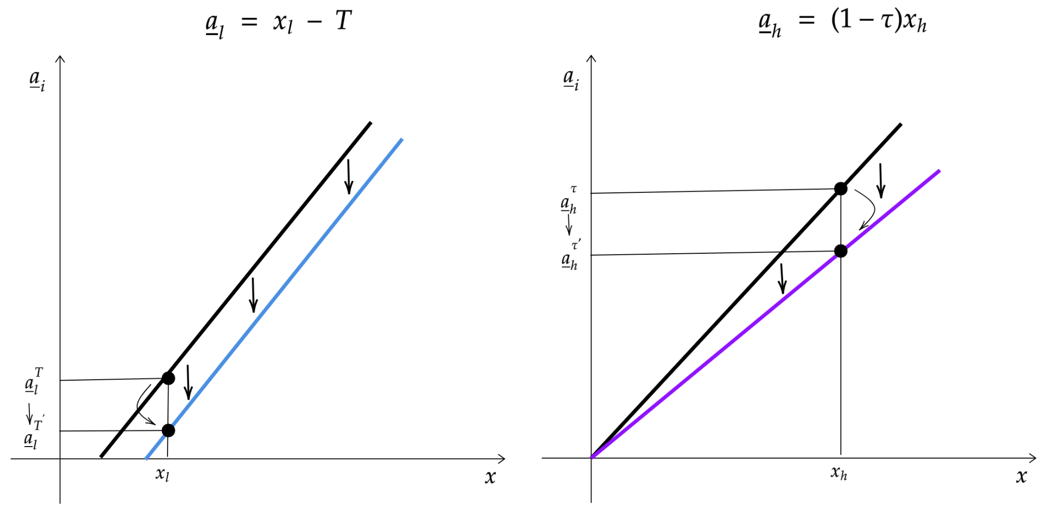
Note: The left figure shows the direct effects of higher progressivity in transfers (while holding indirect distributional effects constant). The right figure shows the direct effects of higher tax progressivity (holding the indirect distributional effects and the width effects fixed).

measured overall progressivity further. In the first case (shown in the left panel), we raise T to $T' > T$, and in the second case (shown in the right panel), we increase τ to $\tau' > \tau$.

In the first case, the direct impact of such reform is a lower $\underline{a}_l$ (and no change in $\underline{a}_h$). There are more marginal workers with lower a, represented by a greater height of the z-perturbed area (the shaded regions in Figure A1). This raises the elasticity of the low type with respect to changes in z.

The second case of more progressive taxation involves two effects. Clearly, the direct impact is a lower $\underline{a}_h$ (and no change in $\underline{a}_l$), which would tend to increase the elasticity of the high type (due to a greater height of the z-perturbed area). However, there is also another important effect: a change in the slope of the threshold line in Figure A2. This slope change could make a difference because the aggregate shifter z varies continuously and determines the width of the perturbed area. In other words, as the return to work $(1 - \tau_i)zx_i$ changes less strongly with a higher τ, the asset threshold responds less strongly with respect to z. Because this second effect (the narrowing width of the z-perturbed regions) works against the former effect (the increasing height of the z-perturbed regions), it is not possible to argue for the presence of a clear-cut direct effect of tax progressivity in the simple model. Note that this does not happen in the case of transfer progressivity because the slope there does not change with respect to T.

Figure A2: Effects of raising the overall progressivity using transfers or taxes



Note: The left figure shows the effect on $\underline{a}_l$ of higher progressivity in transfers by raising T. The right figure shows the effect on $\underline{a}_h$ of higher tax progressivity by raising τ.

C Proofs in Section 2

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

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

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

$$
\begin{aligned}
\frac{\partial N_l}{\partial z} &= \exp(-\underline{a}_l)(1-\lambda), \\
\frac{\partial N_h}{\partial z} &= \exp(-\underline{a}_h)(1+\lambda).
\end{aligned}
$$

we have

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

Also, note that

$$
\begin{aligned}
\frac{\partial N_l}{\partial \omega} &= -\exp(-\underline{a}_l)T\lambda < 0 \\
\frac{\partial N_h}{\partial \omega} &= \exp(-\underline{a}_l)T\lambda > 0.
\end{aligned}
$$

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

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

where

$$\begin{aligned}\underline{a}_l &= z(1-\lambda)-T-T\omega\lambda\\ \underline{a}_h &= z(1+\lambda)-T+T\omega\lambda.\end{aligned}$$

Then, we have

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

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

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

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

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

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

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

Applying this inequality, we have

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

if and only if

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

which is always satisfied.

Proof of Proposition 4 Note that

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

Therefore, we have

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

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

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

we have

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

D Representative-agent (RA) model

We first describe the environment of Model (RA). At the beginning of each period, the stand-in household holds the assets of that period k. The aggregate state variables are the aggregate capital K and the aggregate TFP shock z_k, with the latter following the same stochastic process as in the baseline model. Taking the real wage rate $w(K, z_k)$, the real interest rate $r(K, z_k)$, and 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:

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

$$
\begin{aligned}
\text{subject to} \quad & c + k' \leq (1-\tau_l)w(K, z_k)n + (1 + r(K, z_k))k + T \\
& K' = \Gamma(K, z_k)
\end{aligned}
$$

The household maximizes utility by choosing its optimal consumption c, the next period's capital k', and its labor supply n. The utility of the stand-in household is linear with respect to employment n due to the aggregation theory of 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 $(1-\tau_l)w(K, z_k)n$, current capital k, capital income $r(K, z_k)k$

and government transfers T.

Government then collects taxes on labor earnings $\tau_l wn$ to finance transfers T and government spending 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 presented in (12) and (13).

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_l, 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_l, G, T, the value function $V(k, K, z)$ solves the household's decision problem, and the associated decision rules are

$$\begin{aligned} k'^{*} &= g_k(k, K, z_k) \\ n^{*} &= g_n(k, K, z_k). \end{aligned}$$

2. Prices $r(K, z_k)$, $w(K, z_k)$ are competitively determined following (12) and (13).

3. Government balances its budget:

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

4. Consistency is satisfied: for all K,

$$\begin{aligned} K' = \Gamma(K, z_k) &= g_k(K, K, z_k) \\ L(K, z_k) &= g_n(K, K, z_k). \end{aligned}$$

It is straightforward to calibrate the parameters of Model (RA) using the steady state equilibrium equations. First, β is directly obtained by:

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

Then, given the targets of $T/Y = 0.044$, $L = 0.782$ and $\tau_l = 0.1111$, B is obtained by

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

where

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

Finally, since $Y/K = (K/L)^{\alpha-1}$, we can obtain K/L. This in turn gives us K, and thus Y. We can then obtain T using the calibration target ratio $T/Y = .044$. The resulting calibrated values are $\beta = 0.9901$, $B = 1.0164$, and $T = 0.1277$.

E Heterogeneous-agent models without labor supply indivisibility

Indivisible labor supply is a key feature of our analysis. We illustrate this point by considering a heterogeneous-agent model with divisible labor. The economic environment in this model is mostly identical to the heterogeneous-agent models in the main text, and includes features such as idiosyncratic shocks, progressive taxation, and firm technology. However, one exception is that households can adjust their hours in a fully flexible way under the following period utility function with constant Frisch elasticity γ:

$$U(c,h) = \log c - \xi \frac{h^{1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}. \tag{A2}$$

We consider two different values of $\gamma \in \{1, 2\}$. To illustrate the role of transfers for business cycle fluctuations in this alternative environment, we consider two cases: zero transfers and flat transfers. In the latter case, we target the same moment (4.4% of output) that is used in the main text. For each specification, we also calibrate ξ and β to target the full-time employment rate of 78.2% and the real interest rate of 1% where full-time is defined as hours greater than 0.2.

The results are summarized in Table A1, with two findings in particular being worth highlighting. First, the models without indivisible labor supply have difficulty in generating a sufficiently high volatility of hours worked, echoing the performance of representative-agent real business cycle models (Kydland and Prescott, 1982). Even with a relatively large value of $\gamma = 2$, the volatility of aggregate hours is considerably smaller than in the data. Moreover, these divisible labor models generate average labor productivity that is almost perfectly correlated with output, given that labor supply responses are nearly homogeneous across households and thus do not vary negatively with individual productivity. This is in sharp contrast to our baseline models with labor supply indivisibility (recall Proposition 1 in Section 2). The second notable observation is that the presence of transfers appears almost irrelevant to the cyclicality of average labor productivity, although it does moderately raise the volatility of hours.

Table A1: Results from models without indivisibility

	$\gamma = 1$		$\gamma = 2$		Indivisible	
$T/Y =$	.00	.044	.00	.044	(HA-N)	(HA-F)
σ_Y	1.15	1.16	1.27	1.28	1.48	1.46
σ_C/σ_Y	0.32	0.30	0.31	0.29	0.28	0.27
σ_I/σ_Y	2.77	2.77	2.79	2.79	2.99	2.99
σ_L/σ_Y	0.34	0.35	0.45	0.47	0.64	0.62
σ_H/σ_Y	0.28	0.30	0.36	0.39	0.51	0.60
$\sigma_{Y/H}/\sigma_Y$	0.73	0.70	0.65	0.62	0.54	0.57
$Cor(Y,C)$	0.91	0.91	0.91	0.90	0.85	0.84
$Cor(Y,I)$	0.99	0.99	0.99	0.99	0.99	0.99
$Cor(Y,L)$	0.98	0.98	0.98	0.98	0.96	0.96
$Cor(Y,H)$	0.98	0.97	0.98	0.99	0.95	0.87
$Cor(Y,Y/H)$	1.00	1.00	0.99	0.99	0.95	0.85
$Cor(H,Y/H)$	0.96	0.97	0.96	0.96	0.81	0.81

Note: Each model specification is calibrated to generate the same interest rate and the full-time employment rate.

F Aggregate data

The business cycle statistics are based on the aggregate time-series data from U.S. Bureau of Economic Analysis (BEA), National Income and Product Accounts (NIPA) Tables covering the period from 1961Q1 to 2016Q4. For output, we use the "Real Gross Domestic Product (millions of chained 2012 dollars)" entry in Table 1.1.6. As for consumption, we use expenditures on non-durable goods and services, as reported in Table 2.3.5 (Personal Consumption Expenditure). Investment is constructed as the sum of expenditures on durable goods (Table 2.3.5) and private fixed investments (Table 5.3.5). The real values of consumption and investment are calculated using the price index for Gross Domestic Product from 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 per capita series by dividing the raw data by the quarterly population reported by Cociuba et al. (2018).

A target statistic regarding the size of income-security transfers is based on the aggregate data obtained also from the BEA NIPA Tables. Specifically, we use data from Table 3.12 (Government Social Benefits) on the Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income, Temporary Disability Insurance, and medical care (Medicaid, General Medical Assistance, and state child healthcare programs). Note that we do not include large programs such as Medicare, unemployment insurance, and veterans' benefits.

G Micro data

For the transfer-related 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 and has a monthly survey period. The SIPP covers a wide range of information on income, wealth, and participation in various transfer programs. We choose samples from the first wave to the ninth wave of the SIPP, covering the years 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 all variables at the household level. Data sets in the SIPP contain not only household variables but also individual variables so, in order to generate a household variable from its corresponding individual variable, we take the following steps. First, we identify households by their sample unit identifier (SSUID) and their sample household address identifier (SHHADID). Second, we add up the values of the variable in question for all members of the same household. The government transfers used to infer the degree of progressivity are based on a broad range of transfer programs including Supplemental Security Income (SSI), Temporary Assistant for Needy Family (TANF), the Supplemental Nutrition Assistance Program (SNAP), the Supplemental Nutrition Program for Women, Infants, and Children (WIC), childcare subsidies and Medicaid. We do not include age-dependent programs such as Social Security and Medicare. We also construct a broad household income variable: it consists of labor income, income from financial investments, and property income. We consider households whose head is aged between 23 and 65, and the results we presented are almost the same as for alternative age ranges around these limits. Finally, we convert the nominal values of all these variables to 2001 U.S. dollars using the CPI-U.

The empirical analysis in Section 6 is based on the PSID data. We choose samples for the period of 1969–2010. To avoid the oversampling of low-income household heads, we exclude households listed in the Survey of Economic Opportunity. We also drop the samples whose wage is below one half of the minimum wage. The nominal values are again converted into 2001 U.S.

dollars using the CPI-U.

H Estimation of idiosyncratic productivity risk

We estimate the persistence of idiosyncratic productivity risk in the U.S. using the PSID data, following Heathcote, Storesletten and Violante (2010). Our measure of productivity is defined as a worker's hourly wage relative to other individuals. We consider household heads between the ages of 18 and 70, and whose wages were observed for at least four consecutive periods.[31] To focus on full-time workers, we drop the samples whose annual hours worked was less than $1,000$.

We run the ordinary least squares regression on the logarithm 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 variable that contains a wide range of individual abilities valued by the labor market. This stochastic process is composed of the summation of a persistent process $\eta_{i,j}$ and a transitory process $\nu_{i,j}$ as described by:

$$\begin{aligned} x_{i,j} &= \eta_{i,j} + \nu_{i,j}, \nu_{i,j} \sim N(0, \sigma_\nu^2), \\ \eta_{i,j}^{'} &= \rho_\eta \eta_{i,j-1} + \epsilon_{i,j}^{'}, \epsilon_{i,j}^{'} \sim N(0, \sigma_\epsilon^2). \end{aligned} \tag{A3}$$

We use a minimum distance estimator to estimate the parameters of the process. This method is used to find parameters that minimize the distance between the empirical and theoretical moments. We take the covariance matrix of the residual $x_{i,j}$ as our moments, and denote θ by the vector $(\rho_\eta, \sigma_v, \sigma_\epsilon)$. We then let $m_{j,j+n}(\theta)$ be the covariance of the labor productivity between age j and $j+n$ individuals, and define $\hat{m}_{j,j+n}$ as the empirical counterpart of $m_{j,j+n}(\theta)$. We use the following moment conditions:

$$E\left[\hat{m}_{j,j+n} - m_{j,j+n}(\theta)\right] = 0 \tag{A4}$$

where

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

[31] We use a somewhat less restricted age range in order to obtain a large number of samples. Note that we impose stricter restrictions on wages and hours, which would naturally remove irrelevant samples such as retirees. Thus, a change in the age band leads to only relatively small changes in the estimated persistence of idiosyncratic shocks.

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.[32]

I Numerical methods used for the heterogeneous-agent models

I.1 Solving for the equilibrium with aggregate risk

The models which aggregate risk are solved with 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 those individual policy functions (*the outer loop*). We iterate the two steps until the forecasting rules converge—that is, when the difference between the old forecasting rule used in the inner loop and the new forecasting rule generated in the outer loop becomes small enough.

I.1.1 Inner loop

In the inner loop, we solve for the following value functions: $V(a, x_i, K, z_k)$, $V^E(a, x_i, K, z_k)$ and $V^N(a, x_i, K, z_k)$. These value functions are stored on a non-evenly spaced grid for a and an evenly-spaced grid for K, with the number of grid points being $n_a = 400$ and $n_K = 40$, respectively. Unlike Chang and Kim (2006, 2007) and Takahashi (2014), we discretize the stochastic processes for x_i and z_k by using the Rouwenhorst (1995) method. We find that the approximation of continuous AR(1) processes with our estimate featuring very high persistence is considerably

[32] 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.

better with the Rouwenhorst method given the same number of grid points.[33] Our baseline results are based on $n_x = 10$ and $n_z = 5$, both of which replicate the true parameters of the continuous AR(1) processes very precisely.

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 solve the following problems

$$V^E(a, x_i, K, z_k) = \max_{\substack{a' \geq \underline{a}, \\ c \geq 0}} \left\{ \log c - B\bar{n} + \beta \sum_{j=1}^{N_x} \pi^x_{ij} \sum_{l=1}^{N_z} \pi^z_{kl} V(a', x'_j, \hat{K}', z'_l) \right\} \tag{A5}$$

subject to

$$c + a' \leq \tau(e, \bar{e}) e(\hat{w}(K, z_k)) + (1 + \hat{r}(K, z_k))a + T(\hat{w}(K, z_k), \hat{r}(K, z_k))$$

and

$$V^N(a, x_i, K, z_k) = \max_{\substack{a' > \underline{a}, \\ c > 0}} \left\{ \log c + \beta \sum_{j=1}^{N_x} \pi^x_{ij} \sum_{l=1}^{N_z} \pi^z_{kl} V(a', x'_j, \hat{K}', z'_l) \right\} \tag{A6}$$

$$c + a' \leq (1 + \hat{r}(K, z_k))a + T(\hat{r}(K, z_k))$$

To evaluate the functional value of the expected value function on $(a', \hat{K}')$ which are not on the grid points, we use the piecewise-linear interpolation. By solving these problems, we obtain the individual policy function for work $g_n(a, x_i, K, z_k)$ by comparing $V^E(a, x_i, K, z_k)$ with $V^N(a, x_i, K, z_k)$. We also obtain conditional policy functions for the optimal $a' : g^E_a(a, x_i, K, z_k)$ as the maximizer of the problem (A5) and $g^N_a(a, x_i, K, z_k)$ as the maximizer of the problem (A6).

I.1.2 Outer loop

In the outer loop, we simulate the model economy based on the information obtained in the inner loop. We note that a key step is to find the market-clearing prices in each period during the simulation. Although this is computationally burdensome, we find that the results without the market-clearing step are substantially misleading, as is consistent with Takahashi (2014) and Chang and Kim (2014).

The measure of households $\mu(a, x_i)$ is approximated by a non-evenly spaced grid on a that is finer than that used in the inner loop (Rios-Rull, 1999) and has 4,000 grid points. The variable K is then constructed by aggregating individual asset holdings over the measure of households: $\int_a \sum_{i=1}^{N_x} a\mu(da, x_i)$. Following Takahashi (2014), we use a bisection method to obtain

[33] Specifically, we use the simulated data from the methods of Rouwenhorst and Tauchen, and estimate the persistence and the standard deviation of the error terms in the AR(1) processes for both aggregate productivity shocks and idiosyncratic shocks (results available upon request).

the equilibrium factor prices in each simulation period 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 (12) and (13).

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 $L^s(K, z_k)$. Specifically, we solve (A5) and (A6) 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 (18) 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

$$n = g_n(a, x_i, K, z_k) = \begin{cases} \bar{n} & \text{if } V^E(a, x_i, K, z_k) > V^N(a, x_i, K, z_k), \\ 0 & \text{otherwise.} \end{cases}$$

By 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

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

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

$$\log K' = a_0 + a_1 \log K + a_2 \log z, \tag{A7}$$

$$\log w = b_0 + b_1 \log K + a_2 \log z, \tag{A8}$$

are updated by ordinary least squares with the simulated sequence of $\{K', w, K, z\}$. Our parametric assumptions regarding the forecasting rules are the same as those made in Chang and Kim (2007, 2014) and Takahashi (2014, 2020). We repeat the whole procedure for the inner and outer loops until the coefficients in the forecasting rules converge.

As is clear in the forecasting rules (A7) and (A8), households predict prices and the future distributions of capital based only on the mean capital stock instead of the entire distribution. Therefore, it is important to check whether the equilibrium forecast rules are precise or not. We summarize the results regarding the accuracy of the forecasting rules for the future mean capital stock K' and for the wage w in Table A2. It is clear that all R^2 values are very high in all specifications. We also check the accuracy statistic proposed by Den Haan (2010). Since our dependent variables are logarithmic, 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% for all cases) and the maximum errors are also reasonably small (not exceeding 0.8% for all cases).

Table A2: Estimates and accuracy of forecasting rules

Model	Dependent variable	Coefficient Const.	$\log K$	$\log z$	R^2	Den Haan (2010) error Mean (%)	Max (%)
(HA-T)	$\log K'$	0.1193	0.9554	0.0940	0.99997	0.088	0.445
	$\log w$	-0.2689	0.4242	0.8037	0.99818	0.086	0.749
(HA-N)	$\log K'$	0.1528	0.9413	0.1170	0.99998	0.087	0.499
	$\log w$	-0.5117	0.5291	0.6683	0.99918	0.060	0.452
(HA-F)	$\log K'$	0.1489	0.9431	0.1154	0.99998	0.086	0.427
	$\log w$	-0.4557	0.5045	0.6826	0.99920	0.058	0.425

I.2 Impulse response functions

There is no generally accepted way to calculate conditional impulse responses in nonlinear models. To compute impulse response functions in this paper, we follow the simulation-based procedure developed by Koop et al. (1996) (see also Bloom et al., 2018):

- Draw $i = 1, ..., N_{sim}$ sets of exogenous random variables for aggregate TFP shocks, each of which have $t = 1, ..., T_{sim}$ periods.[34]

- For each set of i, simulate two sequences, one is from the shock economy and the other is from the no-shock economy.

  1. In the shock economy, simulate all interested variables X_{it}^{shock} for $t = 1, ..., T_{shock} - 1$ as normal (as we do in the outer loop). Then, in period T_{shock}, impose a disturbance on aggregate TFP so that it takes an extreme value (e.g., the lowest one z_1). Simulate the economy as normal for the rest of the periods $t = T_{shock} + 1, ..., T_{sim}$.[35]
  2. In the no-shock economy, simulate all interested variables $X_{it}^{noshock}$ for all the periods without any restrictions. The two economies are different only in terms of the imposition of the extreme shock in period T_{shock}.

- The effect of the disturbance on X is given by the average percentage (or percentage point) difference between the two sequences:

$$\hat{X}_t = 100 \times \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \log\left(X_{it}^{shock}/X_{it}^{noshock}\right) \quad \text{(percentage difference)}$$

$$\hat{X}_t = 100 \times \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \left(X_{it}^{shock} - X_{it}^{noshock}\right) \quad \text{(percentage point difference)}$$

The results are based on $N_{sim} = 2,000$ simulations with each simulation having two sequences of the variables of interest for $T_{sim} = 150$ periods. The responses are equal to zero before T_{shock} by construction. The disturbance then hits the economy at period $T_{shock} = 50$, which we label as the first period in our figures.

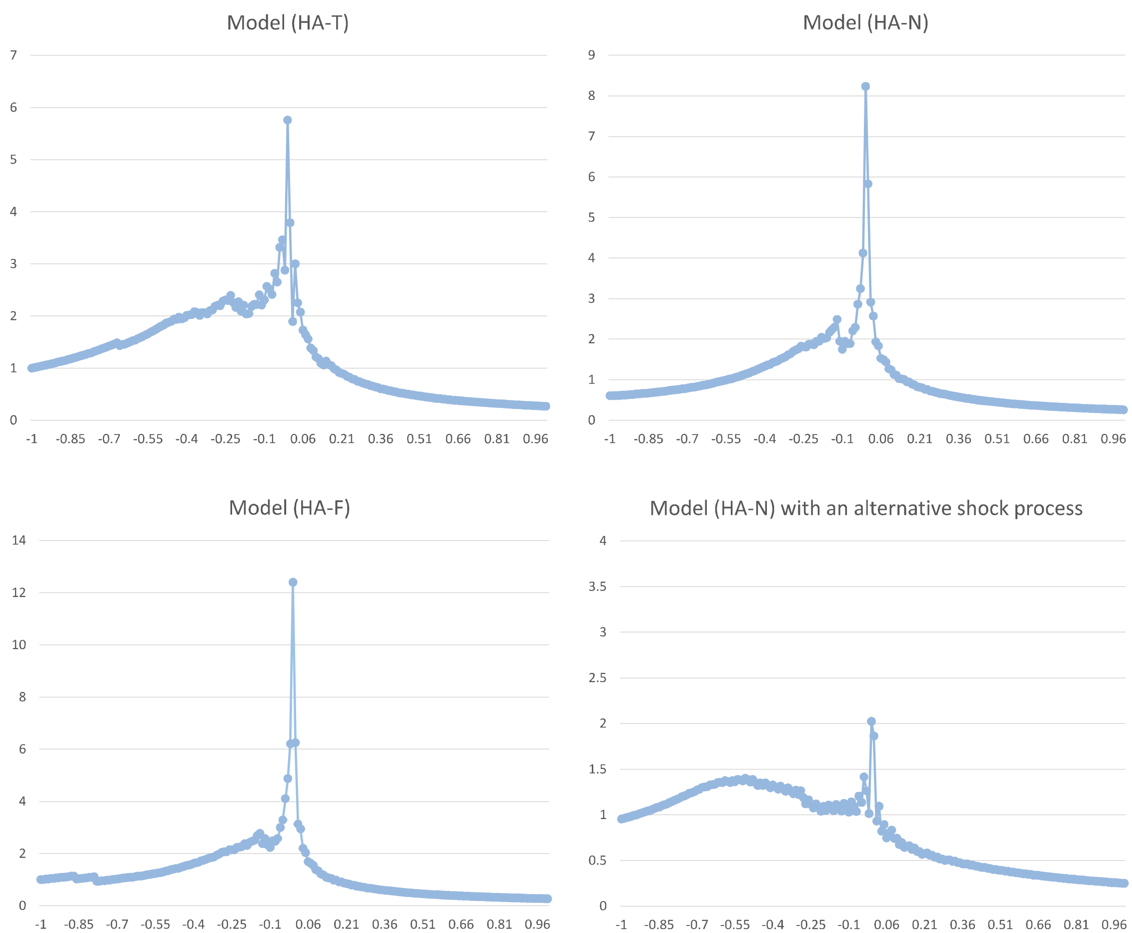
J Additional model results

Table A3 reports business cycle results for several alternative models recalibrated to match the same target statistics as in Table 1. First, we replace the progressive taxation system in (8) with a linear taxation system while keeping the average tax constant. This is helpful for understanding how important the presence of progressive taxation is for business cycles while controlling for

[34] We use a random sampling with Markov chains. That is, by taking as given the index for today's aggregate productivity i and the conditional distribution for tomorrow's productivity $\{\pi_{ij}^z\}_{j=1}^{N_z}$ (i.e., the i-th row of the Markov chain), we draw a random variable $u \sim U[0,1]$ to pick up tomorrow's shock index j. We do so by choosing the highest j satisfying the condition $u < \sum_{k=1}^{j} \pi_{ik}^z$.

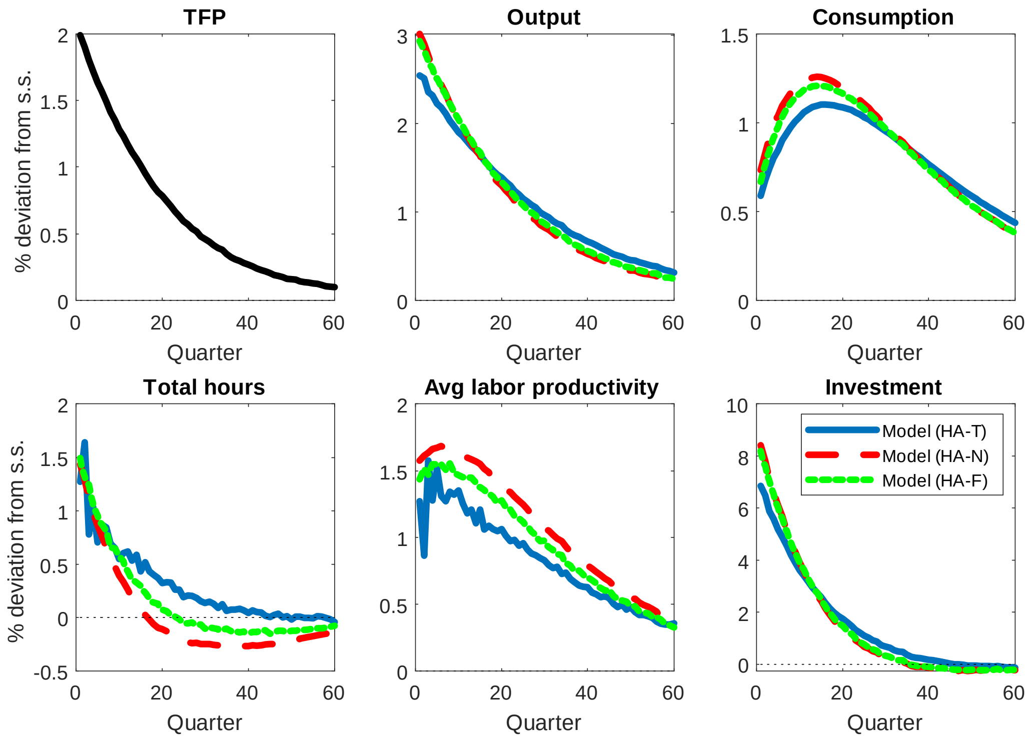
[35] Note that the effect of the disturbance is *persistent* because we sample aggregate productivity using the conditional distribution of the Markov chain.

Figure A3: Arc elasticities for different model specifications



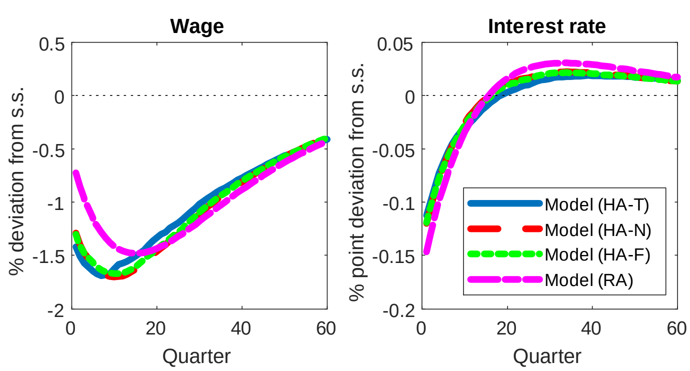
Note: The arc elasticities are computed, based on the reservation raise distribution (Mui and Schoefer 2020). In essence, the reservation raise value of x represents the percentage change in the agent's potential wage that would make the agent to be indifferent between working and non-working. The bottom right panel is from a version of Model (HA-N) with $\rho_x = 0.929$ and $\sigma_x = 0.227$ (Chang and Kim, 2007).

Figure A4: Impulse responses of macroeconomic aggregates with respect to positive TFP shocks



Note: TFP denotes total factor productivity. The figures display the IRFs of macroeconomic aggregates to a positive 2 percent TFP shock with persistence ρ_z.

Figure A5: Impulse responses of equilibrium prices



Note: The figures display equilibrium market-clearing price responses, w_t and r_t, to a negative 2 percent TFP shock with persistence ρ_z. In heterogeneous-agent models, w_t captures the aggregate component of wages conditional on the worker selection in each period.

transfer progressivity. We find that its impact is very minimal for business cycle fluctuations. The second sensitivity check concerns the borrowing limit. The third column in Table 1 reports the results from when we set $\underline{a}$ to zero, and these show that aggregate fluctuations are barely affected by this change. Next, we consider a change in target statistics regarding the variability of idiosyncratic shocks. Recall that the baseline model targets the Gini wage of 0.36. We find that, although its impact is not sizable, a higher wage variation tends to lower the cyclicality of average labor productivity and raise the relative volatility of hours.

Table A3: Sensitivity checks

	Baseline	Linear Taxation	$\underline{a}=0$	Gini wage $=0.35$	Gini wage $=0.37$
σ_Y	1.27	1.23	1.27	1.32	1.24
σ_C/σ_Y	0.27	0.28	0.27	0.26	0.27
σ_I/σ_Y	2.87	2.85	2.86	2.87	2.85
σ_L/σ_Y	0.50	0.47	0.50	0.53	0.48
σ_H/σ_Y	0.73	0.66	0.73	0.80	0.66
$\sigma_{Y/H}/\sigma_Y$	0.64	0.64	0.65	0.62	0.63
$Cor(Y,C)$	0.85	0.87	0.84	0.84	0.85
$Cor(Y,I)$	0.99	0.99	0.99	0.99	0.99
$Cor(Y,L)$	0.92	0.91	0.92	0.94	0.92
$Cor(Y,H)$	0.77	0.78	0.76	0.79	0.79
$Cor(Y,Y/H)$	0.69	0.76	0.68	0.60	0.76
$Cor(H,Y/H)$	0.07	0.18	0.04	-0.02	0.21

Note: Each alternative model is recalibrated to match the same target statistics as in the baseline model.

K Additional empirical results

We provide additional results presented in Section 6 for sensitivity checks.

References

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.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 (3): 1031-1065.

Table A4: Probability of extensive margin adjustment, by wage quintile

	The length of tracking time T								
	5 years			10 years			15 years		
Wage quintile		Switches			Switches			Switches	
in base year	All	Pos only	Neg only	All	Pos only	Neg only	All	Pos only	Neg only
1st	.146	.097	.049	.119	.079	.039	.103	.070	.033
2nd	.093	.060	.033	.080	.052	.029	.072	.046	.026
3rd	.075	.045	.030	.066	.040	.027	.063	.039	.024
4th	.069	.037	.032	.061	.033	.028	.055	.030	.025
5th	.072	.040	.032	.062	.033	.029	.060	.031	.029
Base years	1969–1993 ($J = 25$)			1969–1988 ($J = 20$)			1969–1983 ($J = 15$)		
Avg. no. obs in base years	1,677			1,189			834		
Total no. obs.	41,920			23,783			12,514		
Avg. age	40.2			41.0			41.5		

Note: The full-time employment threshold is set to 1,500 annual hours. Numbers in parentheses show the number of base years. We use samples whose age is between 22 and 64 (inclusive) and who are heads and are not self-employed. "All" refers to the baseline estimates when using both positive and negative switches, whereas "pos only" and "neg only" use only positive ones (i.e., $E_{i,t} = 1$ and $E_{i,t-1} = 0$) and only negative ones (i.e., $E_{i,t} = 0$ and $E_{i,t-1} = 1$), respectively.

Table A5: Full-time 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	-7.3	-10.4	-11.1	-7.1	-8.3	-17.9
2nd	-7.0	-10.5	-10.6	-8.3	-8.9	-16.3
3rd	-5.8	-8.2	-6.3	-7.7	-6.7	-14.9
4th	-4.2	-4.7	-8.0	-7.2	-5.8	-11.1
5th	-1.0	-3.9	-5.2	-3.3	-2.1	-7.4
No. obs.	1,655	1,756	2,007	2,166	2,924	2,802

Note: The full-time employment threshold is set to 1,500 annual hours. The year ranges denote the peak and trough years of each recession. Reported values are percentage changes in the full-time employment rate by wage quintiles (in the peak year of each recession) following the same set of individuals.

Table A6: Full-time 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	-8.5	-3.0	-7.2	-7.0	-9.0
2nd	-4.7	-4.3	-5.7	-6.4	-11.1
3rd	-6.1	-5.2	-6.9	-4.8	-8.7
4th	-4.4	-6.2	-3.9	-5.5	-8.7
5th	-2.5	-6.1	-2.4	-2.7	-5.2
No. obs.	1,547	1,481	1,765	2,454	2,365

Note: The full-time employment threshold is set to 1,500 annual hours. The year ranges denote the peak and trough years of each recession. Reported values are percentage changes in the full-time employment rate by wage quintiles (in the peak year of each recession) following the same set of individuals. Those who experienced unemployment spells in either the peak year or the trough year are excluded. The results for the first recession is omitted because the unemployment information is available only since the 1976 wave (or the year of 1975).

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.

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

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.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2014. "Optimal Tax Progressivity: An Analytical Framework." *The Quarterly Journal of Economics* 132 (4): 1693-1754.

———. 2010. "The Macroeconomic Implications of Rising Wage Inequality in the United States." *Journal of Political Economy* 118 (4): 681-722.

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter. 1996. "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics* 74 (1): 119-147.

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

Mui, Preston and Benjamin Schoefer. 2020. "Reservation Raises: The Aggregate Labor Supply Curve at the Extensive Margin." Unpublished Manuscript.

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

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. 2020. "Time-Varying Wage Risk, Incomplete Markets, and Business Cycles." *Review of Economic Dynamics* 37: 195-213.

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