Aggregate Demand and the Top 1%
Online Appendix

Adrien Auclert Matthew Rognlie

A Discretization of the income process

We assume that log income follows a random walk with negative drift

$$d \log y_{it} = -\mu dt + \sigma dZ_{it} \tag{7}$$

with a reflecting barrier at the lower bound $\underline{y}$. This is known to produce a Pareto stationary distribution with shape parameter $2\mu/\sigma^2$. For tractability, the general equilibrium Huggett-Aiyagari model that we use, adapted from Adrien Auclert and Matthew Rognlie (2016), has a discrete state space for exogenous incomes. Hence, it is necessary to choose some discretization for (7).

To do so, we adapt a simple process from D. G. Champernowne (1953), which produces a discretized Pareto distribution.[6] Assume that log income $x_{it} = \log y_{it}$ can take the values $\{\underline{x} + aj\}$ for $j \geq 0$. Also assume that $x_{it} > \underline{x}$ follows a continuous-time Markov process with a transition rate of u to $x_{it} + a$ and d to $x_{it} - a$, for some $0 < u < d$, with all other transition rates being zero. For $x_{it} = \underline{x}$, assume that the only permissible transition is to $x_{it} + a$, with rate u.

The stationary distribution is Pareto. The stationary distribution of this process is a discretized Pareto distribution. Indeed, if π_j denotes the mass of individuals in state $j \in [0, J]$ in the stationary distribution, stationarity requires that the entering and exiting flows are equalized in each state, i.e.

$$\begin{aligned} u\pi_{j-1} + d\pi_{j+1} &= (u+d)\,\pi_j \qquad j \geq 1 \\ d\pi_1 &= u\pi_0 \end{aligned}$$

[6] For better tractability and a closer approximation to the continuous-time random walk, we start by formulating the process in continuous time, unlike discrete time as in Champernowne (1953), and then derive the implied discrete time transition matrix.

whose solution, enforcing $\sum_{j\geq 0}\pi_j = 1$, is

$$\pi_j = \left(1 - \left(\frac{u}{d}\right)\right)\left(\frac{u}{d}\right)^j$$

and therefore in particular, for any $x_j = \underline{x} + aj$

$$\Pr\left(x_{it} \geq x_j\right) = \left(\frac{u}{d}\right)^j = e^{\left(x_j - \underline{x}\right)\frac{\log\left(\frac{u}{d}\right)}{a}}$$

and hence for any $y_j = e^{x_j}$

$$\Pr\left(y_{it} \geq y_j\right) = \left(\frac{\underline{y}}{y_j}\right)^{\frac{-\log\left(\frac{u}{d}\right)}{a}}$$

We recognize this as the CDF of a discretized Pareto distribution with scale parameter (minimum value) $\underline{y} = e^{\underline{x}}$ and shape parameter

$$\alpha = \frac{-\log\left(\frac{u}{d}\right)}{a} \tag{8}$$

Drift and volatility. Given u and d, both the drift and squared volatility of this process are constant, with

$$\begin{aligned} \mu &= a(d-u) \\ \sigma^2 &= a^2(d+u) \end{aligned}$$

Inverting this relationship, for given μ and σ^2, we have

$$u = \frac{1}{2}\left(\frac{\sigma^2}{a^2} - \frac{\mu}{a}\right) \tag{9}$$

$$d = \frac{1}{2}\left(\frac{\sigma^2}{a^2} + \frac{\mu}{a}\right) \tag{10}$$

Plugging into (8) and simplifying gives

$$\alpha = \frac{1}{a}\log\left(\frac{1 - \frac{a\mu}{\sigma^2}}{1 + \frac{a\mu}{\sigma^2}}\right) \tag{11}$$

Note that in the limit $a \to 0$, $\log\left(1 \pm \frac{a\mu}{\sigma^2}\right) \sim \frac{a\mu}{\sigma^2}$, so that (11) reduces to $\alpha = \frac{2\mu}{\sigma^2}$. This is exactly the formula for α in (4). Hence, as the discretization becomes finer, the relationship between μ, σ, and α approaches that of our idealized income process, the geometric random walk with negative drift and a lower reflecting barrier.

Calibrating the process. Given that $\underline{x}$ has to be chosen to achieve our normalization $\mathbb{E}[y_{it}] = 1$, our process has three free parameters (a, u, d).

The Lorenz curve $L(u)$ of a Pareto with shape α is $1 - L(u) = (1-u)^{1-\frac{1}{\alpha}}$. As mentioned in the text, given a value for the top 1% share, we can then back out the implied Pareto α using

$$\alpha = \frac{1}{1 - \frac{\log(\text{top } 1\% \text{ share})}{\log(1\%)}}$$

Calibrating α on the basis of the top 1% share in 1980, (8) then provides one restriction on (a, u, d). Our calibration of $\sigma^2_{1980} = 0.02$ provides another. Given a, these jointly pin down u and d, which in turn imply $\mu = a(d-u)$, by

$$\begin{aligned} \frac{u}{d} &= e^{-a\alpha} \\ u + d &= \frac{\sigma^2}{a^2} \end{aligned}$$

The remaining choice is a. This is made primarily on computational grounds. Since we require finitely many states for computation, it is necessary to truncate the set $\{\underline{x} + aj\}$ at some maximum $\underline{x} + aJ$. To avoid truncation bias, we pick aJ high enough such that only 0.001% of aggregate income is earned at or above this state in the ideal Pareto distribution for our initial calibration, writing $e^{-(\alpha-1)aJ} = 10^{-5}$. It follows that $aJ = 7.83$, so that the maximum income state is approximately 2500 times higher than the minimum income state. Since the algorithm in AR is $O(J^2)$, we set $J = 40$ for reasonable computation time, implying $a \approx 0.2$.

To map the process to discrete time, we take the matrix exponential to convert the transition *rate* matrix Σ to a Markov transition matrix Π: $\Pi = e^{\Sigma}$.

Changing the distribution. We consider a decline in α from 2.48 to 1.92 to match the rise in the income share of the top 1% since 1980. Using (11), there are various combinations of changes in (a, μ, σ) that can replicate this decline. To ensure that the truncation stays accurate, we vary $a \propto \alpha^{-1}$, such that the maximum state $\underline{x} + aJ$ remains

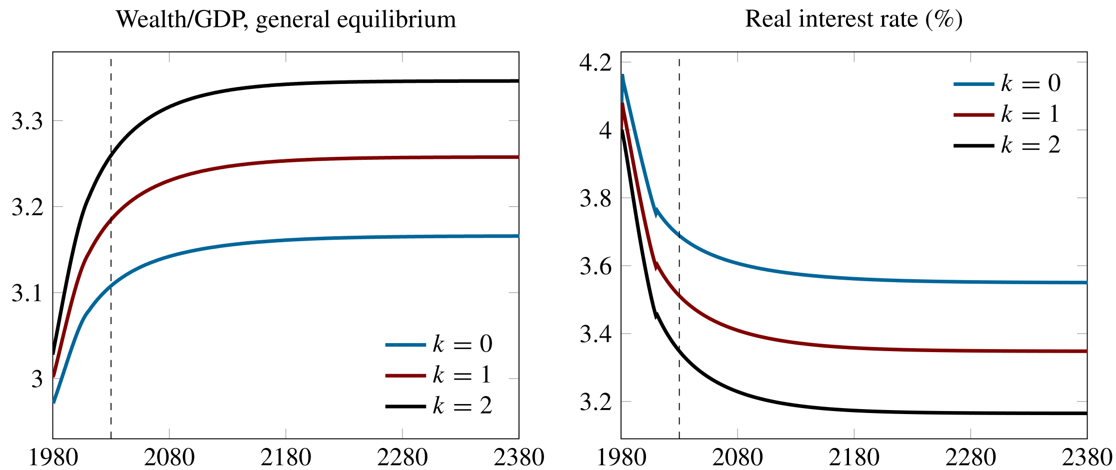

Figure 3: Simulated paths for W_t^{PE} and r_t^{GE}

at the same percentile of the Pareto distribution. (It follows that all other states remain at the same percentiles as well.)

Given $a \propto \alpha^{-1}$, it is clear from (11) that $a\mu/\sigma^2$ must be unchanged. The three experiments in table 1, also discussed in the main text, represent different choices of μ and σ^2 that accomplish this: either (1) $\mu \propto \alpha$ and σ constant ($k = 0$), (2) μ constant and $\sigma^2 \propto \alpha^{-1}$ ($k = 1$), or (3) $\mu \propto \alpha^{-1}$ and $\sigma^2 \propto \alpha^{-2}$ ($k = 2$). The third choice yields unchanged u and d in (9) and (10), and this produces an unchanged transition matrix that is particularly useful for the decomposition (5).

When computing transition dynamics, it is also necessary to specify how the parameters (a, μ, σ) adjust over time to the new steady state. We continue to require that $a\mu/\sigma^2$ is unchanged at all times, so that $a \propto \alpha^{-1}$ in every year. We choose a quadratic trend of a from 1980 to 2011, such that a is consistent with the actual top 1% income share in 1980 and 2011, and minimizes the average square difference between the model and data for the top 1% share across intermediate years. We then assume that σ^2 follows the same convergence path from 1980 and 2011, and infer the implied μ from $\mu \propto \sigma^2/a$. From 2011 onward, we assume all parameters (a, μ, σ) are constant at their new steady-state values.

B Simulated paths for W^{GE} and r^{GE}

Figure 3 plots the perfect foresight transition dynamics of the general equilibrium version of our model (with the benchmark experiment $k = 2$), from its initial steady state to its long-run new steady state after the rise in the Pareto tail of the income distribution. As discussed in the main text and elsewhere in the appendix, we phase in the rise in inequality from 1980 to 2011. The slight discontinuity in interest rates around 2011 on the right panel reflects the end of this phase-in, with income inequality staying constant thereafter, but interest rates continuing to decline.

Note, however, that convergence is much faster than the convergence of partial equilibrium wealth in figure 1: whereas wealth/GDP is only a fraction of the way toward its new steady state in partial equilibrium by 2030 (dashed line) in figure 1, it is a majority of the way toward its new steady state in general equilibrium by 2030 in figure 3.

The long-run increase in wealth/GDP is also much less dramatic in general equilibrium: even when $k = 2$, it increases by slightly over 10% in figure 3, whereas partial equilibrium wealth/GDP in figure 1 nearly doubles. The increase is much smaller in general equilibrium because the endogenous fall in interest rates discourages households from accumulating so much wealth. In our calibration, this margin is more elastic with respect to interest rates than the capital demand margin on the production side of the economy, and therefore bears most of the burden of equilibriation.[7]

C Results for the wealth distribution

Our model also endogenously generates a household wealth distribution. Although it is not the primary focus of this paper, it is also edifying to study this distribution, especially at the top, and how it varies as we make the income distribution more concentrated.

Figure 4 shows the shares of wealth and income held by the top 1% and 0.1% in the model. Several features are apparent:

a) Wealth is endogenously much more concentrated than income. In our 1980 calibration, the top 1% wealth share is 23.6% and the top 0.1% wealth share is 6.5%. (For comparison, these are just slighly below the Emmanuel Saez and Gabriel Zucman (2016) estimates, which are respectively 24.3% and 8.0% for that year.)

[7] The 1980 level of wealth/GDP is differs in the left panel figure 3 for each choice of k. This is because we are displaying perfect foresight paths *after* the shock becomes known. Due to capital adjustment costs, the anticipated general equilibrium adjustment of r leads to changes in the valuation q of capital from its pre-shock steady state. Although r declines in the long run for every k, they initially increase by enough in the $k = 0$ case to offset that, leading to a decline on impact in wealth/GDP from its previous steady-state level of 3. In the $k = 2$ case, by contrast, the long-run decline in r dominates, leading to an increase in wealth/GDP on impact.

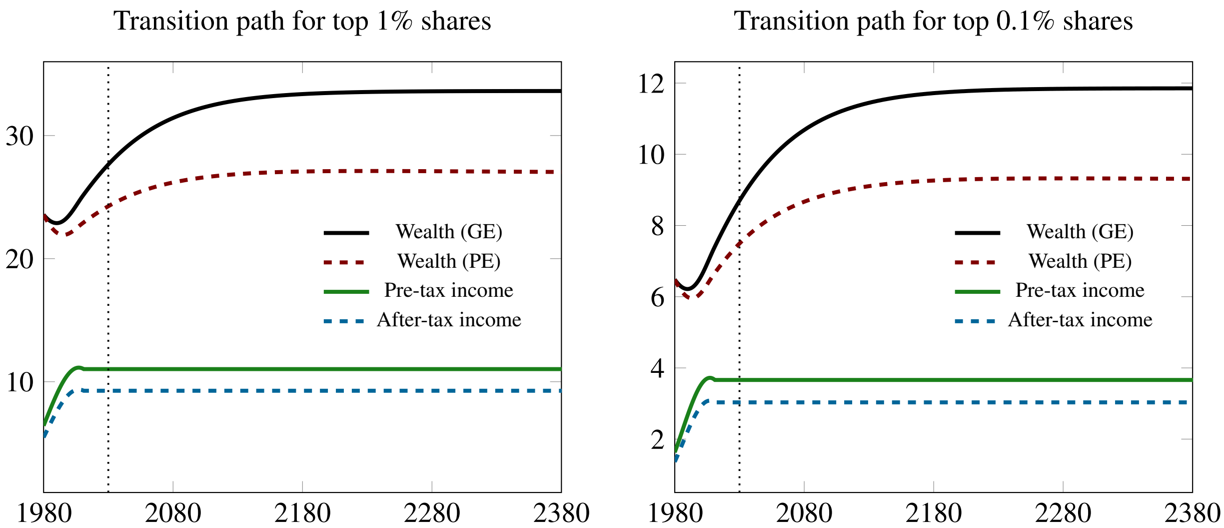

Figure 4: Simulated concentration of wealth and income in the model.

b) A positive, permanent shock to income concentration induces an upward movement in wealth concentration. The induced wealth concentration effect is, in fact, somewhat larger in percentage point terms, although not in relative terms. Following a rise in the top 1% pre-tax income share from 6.4% to 11.0% (1.72x), the top 1% wealth share goes from 23.6% to 33.6% in general equilibrium (1.42x), and from 23.6% to 27.0% (1.14x) in partial equilibrium.

c) On its own, the increase in income inequality is not able to match the trends in wealth concentration documented by Saez and Zucman (2016). Our model first misses the level: their latest (2012) numbers indicate a top 1% share of 41.8%, and a top 0.1% share of 22.0%. By contrast our *steady-state* general equilibrium numbers are, respectively, 33.6% and 11.8%. It also misses the trend: by 2011, when the rise in the top 1% income share is entirely phased in, wealth concentration in our simulations has barely moved at all from its original level. Slightly less than half of the transition occurs in the first 50 years. This delayed convergence has echoes of the slow convergence obtained for the Pareto tail of income in Xavier Gabaix, Jean-Michel Lasry, Pierre-Louis Lions and Benjamin Moll (2016), although here the dynamics arise from the consumption-savings decisions of households rather than from a stochastic income process.

d) Steady-state wealth is more concentrated in our general equilibrium experiments than in our partial equilibrium experiments. This is because interest rates fall in general equilibrium, and, perhaps suprisingly given that lower real interest rates make wealth accumulation more difficult, declining interest rates lead to higher wealth concentration in our model.

Figure 5 shows a different feature of the wealth and income distributions: the tail Pareto parameter. This is inferred from our simulation results by comparing shares held by the top 0.01% and the top 0.001%, and then using the Pareto identity

$$1 - \frac{1}{\alpha} = \frac{\log\left(\frac{0.01\%\text{ share}}{0.001\%\text{ share}}\right)}{\log\left(\frac{0.01\%}{0.001\%}\right)} \tag{12}$$

The comparison of the top 0.01% and 0.001% shares is chosen because the wealth distribution is roughly Pareto this far out in the tail, and because approximation error starts becoming more significant at higher quantiles.

Figure 5 provides numerical confirmation of an analytical result in Jess Benhabib, Alberto Bisin and Shenghao Zhu (2015), which finds that in the stationary steady state of standard Bewley model without idiosyncratic return risk, the wealth distribution has a Pareto tail with the same parameter as the income distribution. The figure also confirms that convergence to a thicker tail for wealth occurs more slowly than for income.

Another important lesson emerges from the contrast between figures 4 and 5. In figure 4, wealth appears substantially more concentrated than income, even though in figure 5 both distributions have the same steady-state Pareto tail parameters. In short, a larger fraction of wealth than income is held by the top 1%, but in the model the shape of the distribution *within* the top 1% (and, even more so, higher percentiles) is roughly the same for wealth and income.

Taken as a whole, these results suggest that models in the S. Rao Aiyagari (1994) tradition may be able to match a substantial component of wealth inequality when augmented with income processes that generate Pareto tails. They are unable, however, to match phenomena that are specific to the right tail of the wealth distribution. These include the Pareto shape parameter of the tail, which Saez and Zucman (2016) indicates to be 1.43 in 2011.[8] The model here also has difficulty matching recent increases in wealth inequality that have been largely confined to the tail, also documented in Saez and Zucman (2016). To obtain these features of the wealth distribution, it is likely necessary to add additional elements to the model—for instance, idiosyncratic return

[8]This is inferred analogously to (12) by comparing the 0.1% and 0.01% shares for that year, which were 20.3% and 10.1%, respectively.

Transition path for implied tail Pareto parameters

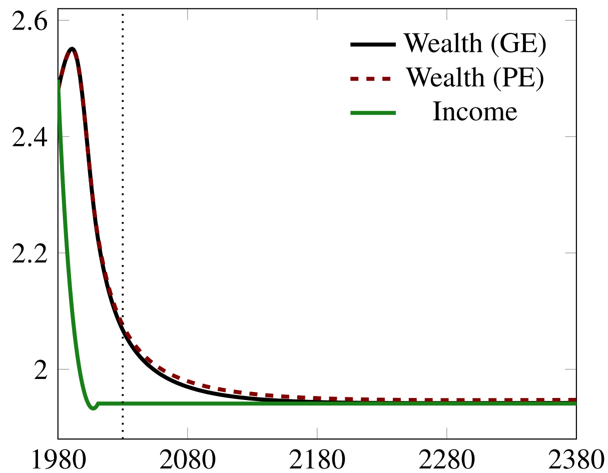

Figure 5: Simulated dynamics of tail thickness in the model.

risk, entrepreneurship, or bequests. Benhabib, Bisin and Zhu (2015) and Mariacristina De Nardi and Giulio Fella (2017) outline some of the possibilities.

D Obtaining our decompositions (5) and (6)

Recall that our discretization of the income process implies a $J \times J$ Markov transition matrix Π, together with levels of incomes $y_1, \ldots y_J$ where $y_j = e^{\underline{y}+aj}$, and that in our $k = 2$ experiment the transition matrix Π is fixed as we change the income levels y_j are changed.

Write $W(\Theta, \Pi, y_1, \ldots y_J)$ for the steady-state level of wealth generated by our partial-equilibrium household model, given parameters $\Theta = (\beta, \nu, r, \tau_r)$, together with Π and y_j. Given fixed Θ, Π, a first-order Taylor expansion of W yields a total change in W equal to

$$dW = \sum_{j=1}^{J} \frac{\partial W_{y_j}}{\partial y_j} dy_j + o\left(\|dy\|\right) \tag{13}$$

Write π_j for the weight of y_j in the stationary distribution induced by Π, then (13) implies that

$$\frac{dW}{W} = \sum_{j=1}^{J} \pi_j \frac{\partial W_{y_j} / \partial y_j}{\pi_j W} dy_j + o\left(\|dy\|\right)$$

Write $\epsilon_{Wj} \equiv \frac{\partial W_{y_j}/\partial y_j}{\pi_j W}$, and drop higher-order terms for notational simplicity. This yields

$$\frac{dW}{W} = \mathbb{E}\left[\epsilon_{Wj} dy_j\right] = \text{Cov}\left(\epsilon_{Wj}, dy_j\right) + \mathbb{E}\left[\epsilon_{Wj}\right]\mathbb{E}\left[dy_j\right]$$

But notice that by construction our income process $\mathbb{E}\left[dy_j\right] = 0$, so we obtain (5).

Similarly, consider the change in consumption at date 0 induced by a change in income at date 0 alone. We can write date-0 aggregate consumption as $C_0\left(\Theta, \Pi, \mathbf{y}; y_{01}, \ldots y_{0J}\right)$, where now $\mathbf{y} = \left(y_1, \ldots y_J\right)$ represents income in each state at all future dates and is held fixed. Then a first-order Taylor expansion of C_0 with respect to this date-0 change is

$$dC_0 = \sum_{j=1}^{J} \pi_j \frac{\partial C_{y_{0j}}}{\pi_j \partial y_{0j}} dy_{0j} + o\left(\|dy_0\|\right)$$

But note that

$$C_0 = \sum_{j=1}^{J} \pi_j \int c_0\left(b, y_j\right) d\Psi_j\left(b\right)$$

where $c_0\left(b, y_j\right)$ is the date-0 policy function of agents with wealth level b and income level j, and $\Psi_j\left(b\right)$ is the density function for wealth b conditional on income being y_j. Hence $\frac{\partial C_{y_{0j}}}{\pi_j \partial y_{0j}} = \int MPC_j\left(b\right) d\Psi_j\left(b\right)$, the average marginal propensity to consume of agents with income level j, which we simply write MPC_j. This delivers

$$dC_0 = \mathbb{E}\left[MPC_j dy_{0j}\right]$$

and noting once again that $\mathbb{E}\left[dy_{0j}\right] = 0$, we obtain (6).

References

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

Auclert, Adrien, and Matthew Rognlie. 2016. "Inequality and Aggregate Demand." Manuscript.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2015. "The Wealth Distribution in Bewley Economies with Capital Income Risk." *Journal of Economic Theory*, 159, Part A: 489–515.

Champernowne, D. G. 1953. "A Model of Income Distribution." *The Economic Journal*, 63(250): 318–351.

De Nardi, Mariacristina, and Giulio Fella. 2017. "Saving and Wealth Inequality." *CEPR Discussion Paper No 11746*.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2016. "The Dynamics of Inequality." *Econometrica*, 84(6): 2071–2111.

Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." *The Quarterly Journal of Economics*, 131(2): 519–578.