THE ECONOMIC JOURNAL

The Economic Journal, **132** (*April*), 906–925 https://doi.org/10.1093/ej/ueab095
Advance Access Publication Date: 4 December 2021

HETEROGENEOUS DYNASTIES AND LONG-RUN MOBILITY*

Jess Benhabib, Alberto Bisin and Ricardo T. Fernholz

Recent empirical work has demonstrated a positive correlation between grandparent-child wealth rank, even after controlling for parent-child wealth rank, and a positive correlation between dynastic wealth ranks across almost six hundred years. We show that a simple heterogeneous agents model with idiosyncratic wealth returns generates a realistic wealth distribution, but fails to capture these long-run patterns of wealth mobility. An auto-correlated returns specification of this model also fails to capture both short- and long-run mobility. However, an extension of the heterogeneous agents model that includes permanent heterogeneity in wealth returns is able to simultaneously match the wealth distribution and short- and long-run wealth mobility.

Recent heterogeneous agents modelling of consumption-saving decisions has successfully been able to identify the main drivers of wealth inequality (Quadrini, 2000; Castañeda *et al.*, 2003; Cagetti and De Nardi, 2006; Krusell and Smith, 2015; Hubmer *et al.*, 2016; Benhabib *et al.*, 2019; Kindermann and Krueger, 2021). This literature shows that heterogeneous agents models with stochastic labour earnings and idiosyncratic returns to wealth can produce fat-tailed distributions of wealth that match the data well.[1] These models can also fit reasonably well the inter-generational social mobility of wealth, producing a realistic parent-child wealth-rank correlation (Benhabib *et al.*, 2019).

More recently, however, empirical results suggest a significant grandparent-child wealth-rank correlation even after controlling for the effects of parent wealth on child wealth (Boserup *et al.*, 2014). Furthermore, even long-run wealth-rank correlations appear to persist across generations (Clark, 2014; Clark and Cummins, 2015; Barone and Mocetti, 2016). With respect to this dimension of inter-generational mobility, the heterogeneous agents models in the literature do not fare well, in that they cannot generate a large enough coefficient for grandparent-child wealth rank nor a large enough correlation for dynastic wealth ranks over very long time periods. We discuss the theoretical reasons why this class of models produces limited long-run rank-wealth correlations in Subsection 2.1. In Section 3 we confirm this by means of simulation analysis.

In this paper we extend a simple heterogeneous agents model to introduce permanent heterogeneity in the rate of return to wealth across generations. In other words, we allow households in some dynasties to have their wealth grow faster on average than households in other dynasties. This can be seen as a formalisation of a latent factor representation of persistent cultural and institutional factors suggested in the literature in political economy and sociology, along the lines of Bourdieu (1984; 1998), Bisin and Verdier (2001) and Acemoglu and Robinson (2008). In

* Corresponding author: Alberto Bisin, Department of Economics, New York University, 19 W. 4th St., 6th Floor, New York, NY 10012, USA. Email: alberto.bisin@nyu.edu

This paper was received on 3 June 2020 and accepted on 29 November 2021. The Editor was Frederic Vermeulen.

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.5634269.

We thank the editor and referees for their insightful comments.

[1] See Benhabib *et al.* (2017) for a discussion of the relative role of earnings and returns to wealth.

Subsection 2.1 we show theoretically that such a model has the potential to generate a strong inter-generational rank correlation of wealth, also in the long run. In Section 3 we confirm that a calibrated permanently heterogeneous rank-based model can produce both a fat-tailed distribution of wealth that matches the data well as well as strong inter-generational correlations akin to those documented in the data.

It is not difficult to envision other extensions of simple heterogeneous agents models that could produce significant grandparent-child and even long-run wealth-rank correlations, e.g., postulating inter-generational auto-correlation in earnings and in the rate of return to wealth. However, the evidence is not favourable to the existence of independent direct causal effects across generations, beyond parent-child effects. We discuss this evidence in the next section. Extending the model along these lines requires postulating very strong inter-generational auto-correlation in earnings and in the rate of return to wealth to capture long-run persistence, which does not appear plausible and certainly not parsimonious as an explanation. Indeed, in Subsection 3.2.1 we show by means of simulation analysis that the long-run auto-correlations of wealth ranks in the data can be generated in principle by models specifying auto-correlated returns to wealth, but at the cost of excessively high parent-child and grandparent-child wealth-rank correlations with respect to the data.

Finally, we compare the implications of the permanently heterogeneous rank-based model and the auto-correlated returns to wealth model with respect to parent-child return-rank correlation. We show in Subsection 3.2.1 that the model with permanent heterogeneity produces in our calibration a small parent-child correlation of returns, close to that documented by Fagereng *et al.* (2020). The auto-correlated returns to wealth model on the other hand also produces an excessively high parent-child return-rank correlation. We argue that this is suggestive of persistent institutional factors as mechanisms for sustaining long-run wealth persistence, rather than of direct inter-generational mechanisms like cultural transmission.

1. Long-Run Wealth-Rank Correlation

In this section we briefly discuss the evidence documenting wealth-rank correlations across generations and its interpretation in the literature. First of all, a positive correlation between grandparent-child wealth rank, even after controlling for parent-child wealth rank, is documented in Boserup *et al.* (2014), using three generations of Danish wealth data. Since parent and grandparent wealth are correlated, and also possibly measured with error, they implemented a two-stage least squares procedure to identify direct grandparent effects. They found that grandparent effects do not necessarily go through parents and concluded in favour of indirect effects, which they interpreted as ‘social status’. Relatedly, Braun and Stuhler (2018) identified a possible direct causal effect of grandparent-child interactions, exploiting quasi-exogenous variation in the time of grandparents’ death during World War II. They also found no effects of direct contacts between grandparents and grandchildren and concluded in favour of grandparent effects operating through indirect mechanisms. Finally, Warren and Hauser (1997), using data from the Wisconsin Longitudinal Study, found no evidence for an independent influence of grandparents once they conditioned on the status of both parents.

The evidence on long-run dynastic wealth-rank correlation is noteworthy. Clark (2014) and Clark and Cummins (2015) found high persistence of wealth across five generations using data on rare surnames in England and Wales between 1858 and 2012; and Barone and Mocetti (2016) found significant positive wealth elasticities as well as occupational persistence for families in

Florence between 1427 and 2011.[2] These data are necessarily plagued by noise, to the point of being hardly amenable to statistical inference to identify any latent factors responsible for the persistence of wealth, education or occupational status in the long run (Mare, 2011; Braun and Stuhler, 2018). Nonetheless, this documented persistence is consistent with the evidence on grandparent-child correlations, as argued in Stuhler (2012) and Braun and Stuhler (2018). It is also consistent with recent empirical and theoretical studies identifying long-run persistence in cultural traits (see Voth, 2021; Bisin and Moro, 2021 and Bisin and Verdier, 2001 for surveys) and with the evidence of long-run persistence of the effects of institutions, especially of those institutional factors that perpetuate political and economic elites and hence wealth inequality (see Acemoglu and Robinson, 2008; Bisin and Verdier, 2017; 2021 and the work by Pierre Bourdieu, e.g., Bourdieu and Passeron, 1970; Bourdieu, 1984; 1998).

2. Models of Wealth Dynamics

In this section we develop the theory behind our analysis of long-run persistence in rank-wealth correlation. We study rank-based models of wealth dynamics, that is, models in which the growth rate of wealth depends on the wealth rank rather than, e.g., the wealth level. These models are convenient for our analysis as they allow for an analytic characterisation of asymptotic wealth ranks and approximate standard heterogeneous agents models well. In the following, we first introduce a standard rank-based model and relate it to heterogeneous agents models. We then introduce permanent heterogeneity in the rate of return to wealth across generations into the standard rank-based model. Finally, we derive theoretical results about long-run persistence of wealth-rank correlations.

2.1. *Rank-Based Models*

Consider an economy populated by N households, indexed by $i = 1, \ldots, N$. Ranking households by their wealth, let $\rho_t(i)$ denote the wealth rank of household i at time $t \in \mathbb{R}$, so that $\rho_t(i) < \rho_t(j)$ if and only if $w_i(t) > w_j(t)$ or $w_i(t) = w_j(t)$ and $i < j$. We define the ranked wealth processes $w_{(1)} \geq \cdots \geq w_{(N)}$ by $w_{(\rho_t(i))}(t) = w_i(t)$. The aggregate wealth of the economy is then $w(t) = w_1(t) + \cdots + w_N(t)$.

For each household $i = 1, \ldots, N$, wealth dynamics are given by

$$d \log w_i(t) = \alpha_{\rho_t(i)}\, dt + \sigma_{\rho_t(i)}\, dB_i(t), \tag{1}$$

where B_i is a Brownian motion. The parameters α_k and σ_k measure the expectation and the variance of the growth rate of wealth at each rank k.[3] We normalise, without loss of generality, the average growth rate of the economy to zero; that is, $\alpha_1 + \cdots + \alpha_N = 0$. The parameters α_k capture then the expected *relative growth rates* of wealth with respect to the growth rate of the economy. According to Proposition 2.3 of Banner *et al.* (2005), the rank-based model (1) admits a stationary distribution if

$$\alpha_1 + \cdots + \alpha_k < 0 \tag{2}$$

[2] Long-run persistence is also documented by Braun and Stuhler (2018), Lindahl *et al.* (2015), Long and Ferrie (2013) and Modin *et al.* (2013) on occupational and educational attainment, and by Chan and Boliver (2013) and Hertel and Groh-Samberg (2014) on social class.

[3] We acknowledge that this notation implies that high wealth ranks correspond to small rank index k's.

for all $k = 1, \ldots, N-1$. Condition (2) on the parameters α_k suffices to guarantee that no household in the top ranks grows in expectation faster than in the lower ranks, which would cause it to break away from the average population wealth. We show in the next section that this condition is consistent with rates of return to wealth that are constant or even increasing in wealth in a standard heterogeneous agents model of wealth dynamics.

We can now characterise the stationary distribution of the rank-based model.

PROPOSITION 1. *Consider a rank-based model (1) that satisfies (2) and also*

$$\sigma_{k+1}^2 - \sigma_k^2 = \sigma_k^2 - \sigma_{k-1}^2 \tag{3}$$

for all $k = 2, \ldots, N-1$. The ranked wealth processes satisfy

$$\mathbb{E}[\log w_{(k)}(t) - \log w_{(k+1)}(t)] = \frac{\sigma_k^2 + \sigma_{k+1}^2}{-4(\alpha_1 + \cdots + \alpha_k)} \tag{4}$$

for all $k = 1, \ldots, N-1$, where the expectation is taken with respect to the stationary distribution.

It follows then that the expected value of the ratio of wealth in rank k to wealth in rank $k+1$ at the stationary distribution (*i*) is positive for all ranks k;[4] and (*ii*) is increasing in the volatility parameters σ_k, σ_{k+1}. As an illustration, if the expected relative growth parameters α_k were increasing in k and σ_k constant, $\mathbb{E}[\log w_{(k)}(t) - \log w_{(k+1)}(t)]$ would be decreasing in rank until α_k turned positive. We finally note that, by the result in Theorem 2 of Ichiba *et al.* (2011), $w_{(k)}/w_{(k+1)}$ follows a Pareto distribution, with the Pareto parameter for each k depending on the parameters α_k and σ_k according to (4).

2.1.1. *Rank-based model as an approximation*

We introduce a simple heterogeneous agents consumption-saving model, along the lines of Benhabib *et al.* (2011) and Benhabib *et al.* (2019), and show that it can be formally mapped into an approximated rank-based model such as (1). In Section 3 we then show that an appropriate calibration of this model indeed approximates the heterogeneous agents consumption-saving model well.[5]

Consider an economy populated by households who live for one generation, from t to $t+1$, in discrete time. Any household born at time $t \in \mathbb{N}$ has a single child entering the economy at time $t+1$, that is, at its parent's death. Generations of households are linked to form dynasties. A single generation is composed of T subperiods and each household solves a dynamic consumption-savings problem over subperiods, maximising a present discounted Constant Relative Risk Aversion (CRRA) utility function with a joy-of-giving bequest final term (leaving its wealth at death to its child). The household faces idiosyncratic yearly rates of return on wealth $r_{i,t}$ and yearly base earnings $y_{i,t}$ at birth; that is, $r_{i,t}$ and $y_{i,t}$ are stochastic across generations but deterministic inside each generation (more precisely, the rate of return remains constant while earnings grow at a constant growth rate g). All households are ex ante identical, except for the rate of return to wealth, earnings and initial wealth (as bequests). In equilibrium, the inter-generational wealth dynamics for each household $i = 1, \ldots, N$ follow

$$w_i(t+1) = \lambda(r_{i,t})w_i(t) + \beta(r_{i,t}, y_{i,t}), \tag{5}$$

[4] Recall that, by the stationarity condition (2), the sums in the denominator of the right-hand side of (4) are all negative for $k = 1, \ldots, N-1$ where the expression is defined.

[5] More generally, model (1) can be calibrated to approximate many different dynamic models and real-world phenomena that exhibit Pareto-like distributions (Fernholz, 2017).

where $y_{i,t}$ and $r_{i,t}$ respectively denote the yearly base labour income (growing at constant rate g in generation t) and the yearly return on wealth for household i (constant in generation t), and $w_i(t)$ denotes the wealth holdings of household i in generation t. Equation (5) represents wealth accumulation in reduced form, after optimal consumption has been subtracted from the right-hand side. The functions λ and β are obtained as closed-form solutions of the dynamic optimal consumption-saving problem of the household. They are the same for all households $i = 1, \ldots, N$ and respectively represent the inter-generational return on wealth and the present discounted value of labour income, after optimal household consumption, an affine linear function of wealth, has been netted out each period.[6] We refer to model (5) as the standard model.

Let the function $\pi_t(k)$ identify the index i of the kth ranked household at time t, so that $\pi_t(k) = i$ if and only if $\rho_t(i) = k$. The *rank-based approximation* of the standard model (5) is the rank-based model (1) where the parameters α_k and σ_k are defined by

$$\begin{aligned}\alpha_k &= \mathbb{E}[\log(w_{\pi_t(k)}(t+1)/w(t+1)) - \log(w_{\pi_t(k)}(t)/w(t))],\\ \sigma_k^2 &= \mathrm{Var}[\log(w_{\pi_t(k)}(t+1)/w(t+1)) - \log(w_{\pi_t(k)}(t)/w(t))],\end{aligned} \tag{6}$$

for each rank $k = 1, \ldots, N$.[7] At each rank k of the distribution, the parameters α_k and σ_k measure the expectation and the variance of the growth rate of wealth relative to the aggregate for a generation in the standard model (5). Because the rank-based approximation of (5) uses the parameters α_k and σ_k defined by (6), it follows that these parameters represent inter-generational expected relative growth rates and variances in the rank-based model. Like the standard model, then, each new generation in the continuous-time rank-based approximation (1) is born at time $t \in \mathbb{N}$, and these generations are stacked.

The expected relative growth rate parameters α_k represent the main link between the rank-based model (1) and the standard model (5): at the stationary distribution of the standard model, the rank-based expected relative growth rate parameters α_k satisfy, for each rank $k = 1, \ldots, N$,

$$\begin{aligned}\alpha_k &= \mathbb{E}[\log(w_{\pi_t(k)}(t+1)/w(t+1)) - \log(w_{\pi_t(k)}(t)/w(t))]\\ &= \mathbb{E}[\log(w_{\pi_t(k)}(t+1)/w_{\pi_t(k)}(t))]\\ &= \mathbb{E}[\log(\lambda(r_{\pi_t(k),t}) + \beta(r_{\pi_t(k),t}, y_{\pi_t(k),t})/w_{\pi_t(k)}(t))],\end{aligned} \tag{7}$$

since the expected value of aggregate wealth w satisfies $\mathbb{E}[\log w(t+1)] = \mathbb{E}[\log w(t)]$ by stationarity. From (7), we can express the rank-based expected relative growth rate parameters α_k in

[6] More precisely, in Benhabib *et al.* (2011), the functions β and λ depend on (*i*) the generation span T and the growth rate of labour income over time g; (*ii*) preference parameters η, ψ and χ, representing the time discount rate, the elasticity of substitution and the bequest motive, respectively; (*iii*) policy parameters b and ζ, denoting the estate tax on bequests of wealth and the capital income tax rate. They are expressed in closed form as

$$\lambda(r_{i,t}) = (1-b)e^{\bar{r}_{i,t}T}\frac{A(\bar{r}_{i,t})B(b)}{e^{A(\bar{r}_{i,t})T} + A(\bar{r}_{i,t})B(b) - 1},$$

$$\beta(r_{i,t}, y_{i,t}) = (1-b)y_{i,t}\frac{e^{(g-\bar{r}_{i,t})T} - 1}{g - \bar{r}_{i,t}}e^{\bar{r}_{i,t}T}\frac{A(\bar{r}_{i,t})B(b)}{e^{A(\bar{r}_{i,t})T} + A(\bar{r}_{i,t})B(b) - 1},$$

with

$$A(r_{i,t}) = r_{i,t} - \frac{r_{i,t} - \eta}{\psi}, \quad B(b) = \chi^{1/\psi}(1-b)^{(1-\psi)/\psi} \quad \text{and} \quad \bar{r}_{i,t} = (1-\zeta)r_{i,t}.$$

[7] The expectations in (6) are calculated under the stationary distribution of model (5). Note that model (5) has a Brownian motion continuous-time limit; see Saporta and Yao (2005).

terms of β and λ, the parametric functions characterising the solution of the consumption-savings problem underlying the standard model (5):

$$\begin{aligned}\alpha_k &= \mathbb{E}\left[\log\left(\lambda(r_{\pi_t(k),t})\left(1+\frac{\beta(r_{\pi_t(k),t},\, y_{\pi_t(k),t})}{\lambda(r_{\pi_t(k),t})w_{\pi_t(k)}(t)}\right)\right)\right] \\ &= \mathbb{E}[\log(\lambda(r_{\pi_t(k),t}))] + \mathbb{E}\left[\log\left(1+\frac{\beta(r_{\pi_t(k),t},\, y_{\pi_t(k),t})}{\lambda(r_{\pi_t(k),t})w_{\pi_t(k)}(t)}\right)\right]. \end{aligned} \tag{8}$$

Equation (8) provides a simple decomposition of the rank-based expected relative growth rates α_k from (1) in terms of (*i*) the inter-generational return on wealth, adjusted for equilibrium household behaviour, λ, at rank k in the wealth distribution; (*ii*) the present discounted value of labour income, adjusted for equilibrium household behaviour, β, divided by a measure of generational capital income, λw, at rank k in the distribution.

Decomposition (8) is a fundamental interpretation tool in our analysis in that it allows us to map the stability condition for rank-based models in (2) into a condition in terms of β and λ in the standard model (5). In the standard model (5), wealth returns $r_{i,t}$ and labour income $y_{i,t}$ are both independent of the wealth rank of household i at time t. Consequently, the first component of decomposition (8), λ, is independent of wealth rank while the second component of this decomposition, $\beta/(\lambda w)$, is decreasing in wealth rank since wealth w is increasing in rank. In the standard model, then, $\alpha_1 < \alpha_2 < \cdots < \alpha_N$ because the ratio of labour income to capital income is lower at higher ranks in the wealth distribution. A negative relationship between wealth and returns, as in, e.g., models with decreasing returns like Cagetti and De Nardi (2006), is not required to satisfy the stability condition (2). In fact, if follows from this argument that even a positive relationship between wealth and returns in the standard model could be consistent with condition (2).

PROPOSITION 2. *If the standard model (5) is stationary then its rank-based approximation defined by (1) and (6) is also stationary.*

Benhabib *et al.* (2019) present a model of the form (5) with higher returns to wealth at higher wealth ranks and show that this model admits a stationary distribution. Therefore, Proposition 2 implies that the rank-based approximation of this model is also stationary and satisfies the stability condition (2), despite the positive relationship between wealth and returns.

2.2. *Permanently Heterogeneous Rank-Based Model*

We introduce a form of permanent heterogeneity in the expected growth rates of households in the rank-based model (1). For each household $i = 1, \ldots, N$, wealth dynamics are given by

$$d\log w_i(t) = (\gamma_i + \hat{\alpha}_{\rho_t(i)})\,dt + \sigma_{\rho_t(i)}\,dB_i(t) \tag{9}$$

for each household $i = 1, \ldots, N$, with $\gamma_i \in \{\gamma_\ell, \gamma_h\}$ and $\gamma_h > \gamma_\ell$. The parameter γ_i acts as a permanent additive factor to the expectation of the growth rate of wealth: if $\gamma_i = \gamma_\ell$ (respectively $\gamma_i = \gamma_h$), household wealth grows more slowly (respectively quickly) in expectation over time. We assume that n of the households are characterised by $\gamma_i = \gamma_h$, and $N - n$ of the households by $\gamma_i = \gamma_\ell$. We keep normalising the average growth rate of wealth to zero, which in this economy requires $\sum_{k=1}^N \hat{\alpha}_k + \sum_{i=1}^N \gamma_i = \sum_{k=1}^N \hat{\alpha}_k + (N-n)\gamma_\ell + n\gamma_h = 0$. The growth rate of wealth of each household i is nonetheless stochastic, due to the Brownian motion term $\sigma_{\rho_t(i)}\, dB_i(t)$, whose volatility depends on the wealth rank of the household $k = \rho_t(i)$.

To admit a stationary distribution, the permanently heterogeneous model (9) must satisfy a condition that generalises condition (2) for the rank-based model (1) with no heterogeneity. Following Ichiba *et al.* (2011), this condition states that

$$\sum_{k=1}^{m} \hat{\alpha}_k + \tilde{m}\gamma_h + (m - \tilde{m})\gamma_\ell < 0 \quad \text{for all } m = 1, \ldots, N-1;\ \tilde{m} = \min(m, n). \tag{10}$$

Condition (10) ensures that, accounting for the permanent heterogeneity in the expected growth rates of households, no top subset of households grows faster than the aggregate in expectation. This is sufficient to guarantee that the high-growth households (with $\gamma_i = \gamma_h$) in the top ranks do not break away from the rest of the population.

2.3. *Long-Run Wealth-Rank Correlations*

In this section we provide a theoretical characterisation of asymptotic wealth rank for both the standard rank-based model and the model with permanent heterogeneity. We show that permanent heterogeneity is required to generate long-run wealth-rank correlations.

We start with the implications of the rank-based model (1) for mobility. We define occupation times $\xi_{i,k}$ for all i, k, as the fraction of time household i occupies rank k, $\xi_{i,k} = \lim_{T\to\infty}(1/T)\int_0^T 1_{\{\rho_t(i)=k\}}\, dt$. Note that, by definition, the occupation times must add up to one, so that $\sum_{i=1}^{N} \xi_{i,k} = \sum_{k=1}^{N} \xi_{i,k} = 1$. We can now show the following.

PROPOSITION 3. *Occupation times $\xi_{i,k}$ in the standard rank-based model (1) satisfy*

$$\xi_{i,k} = \frac{1}{N} \quad \textit{almost surely (a.s.) for all } i, k. \tag{11}$$

Furthermore, for each household i, the asymptotic wealth rank satisfies

$$\lim_{\tau\to\infty} \mathbb{E}[\rho_{t+\tau}(i)] = \frac{N+1}{2}. \tag{12}$$

This result is a consequence of the fact that all households in model (1) have identical expected wealth dynamics. Therefore, (*i*) they will spend equal time in all ranks, (11); and (*ii*) they must on average approach the same rank asymptotically; hence, necessarily the median of the distribution, (12). In other words, (11)–(12) imply that higher-ranked households today do not occupy in expectation higher ranks in the future as well. As a consequence, the standard rank-based model (1) cannot produce long-run wealth-rank correlations.

This is not the case when permanent heterogeneity is added to the standard rank-based model, as in (9). We turn now to analyse the implications of this model for asymptotic wealth rank. If household i has $\gamma_i = \gamma_\ell$ then, by symmetry, the fraction of time household i spends in each rank k is equal to the fraction of time any other low-growth household spends in each rank k. Thus, we can define the low-growth household occupation times $\xi_{\ell,k}$ such that $\xi_{\ell,k} = \xi_{i,k}$ for all ranks $k = 1, \ldots, N$. Similarly, if we suppose that household j is a high-growth household with $\gamma_j = \gamma_h$ then we can define the high-growth household occupation times $\xi_{h,k}$ such that $\xi_{h,k} = \xi_{j,k}$ for all ranks $k = 1, \ldots, N$. Because the sum of occupation times across all ranks or individual households must equal one, it follows that the low- and high-growth occupation times $\xi_{\ell,k}$ and

$\xi_{h,k}$ must satisfy

$$(N-n)\xi_{\ell,k} + n\xi_{h,k} = 1$$

for all $k = 1, \ldots, N$.

PROPOSITION 4. *Consider a permanently heterogeneous rank-based model (9) that satisfies (3) and (10). Then, the low- and high-growth occupation times $\xi_{\ell,k}$ and $\xi_{h,k}$ satisfy*

$$0 < \xi_{\ell,1} < \xi_{\ell,2} < \cdots < \xi_{\ell,N} < \frac{1}{N-n} \quad a.s.$$

and

$$\frac{1}{n} > \xi_{h,1} > \xi_{h,2} > \cdots > \xi_{h,N} > 0 \quad a.s.$$

Because the occupation times for both low- and high-growth households satisfy $\xi_{i,1} + \cdots + \xi_{i,N} = 1$, Proposition 4 implies that $\xi_{\ell,1} < \xi_{h,1}$ and $\xi_{h,N} < \xi_{\ell,N}$. This means that low-growth households spend more time at the lowest ranks of the wealth distribution across generations than high-growth households. The following theorem uses this result to show that the heterogeneous rank-based model (9) will feature persistence in wealth ranks over infinitely long time horizons.

THEOREM 1. *Consider a permanently heterogeneous rank-based model (9) that satisfies (3) and (10). Then,*

$$\lim_{\tau\to\infty} \mathbb{E}[\rho_{t+\tau}(i)] < \lim_{\tau\to\infty} \mathbb{E}[\rho_{t+\tau}(j)] \quad \textit{if and only if} \quad \rho_t(i) < \rho_t(j) \tag{13}$$

for all households $i, j = 1, \ldots, N$, where the expectations are taken with respect to the stationary distribution.

Theorem 1 implies that the long-run asymptotic household wealth-rank correlation will be positive in the heterogeneous rank-based model. This is because higher-ranked households occupy higher ranks in expectation, due to the underlying persistence heterogeneity (the expectations in (13) are unconditional with respect to whether households i and j are high or low growth).

The intuition for the result in Theorem 1 is worth presenting in some detail as it underlies some of the simulation results in the next section. Because all high-growth households are ex ante identical, the expected asymptotic rank of these households is the median of the top n ranks of the wealth distribution; that is, high-growth households occupy higher ranks in expectation across generations than low-growth households. Similarly, the expected asymptotic rank of low-growth households is the median of the bottom $N-n$ ranks. Without knowing whether a household i is high or low growth, its expected asymptotic rank is thus a weighted average of the medians of the top n and bottom $N-n$ ranks, with the weights equal to the respective probabilities that household i is high growth and that it is low growth. Because higher-ranked households are more likely to be high-growth households, it follows that the weight on the median of the top n ranks is greater for such high-ranked households and hence the expected asymptotic rank is also higher.

To better understand the simulation results in the next section, it is important to emphasise, however, that while the long-run asymptotic household wealth-rank correlation depends on the permanent heterogeneity, the parent-child correlation of the growth rate of wealth is affected negatively in a crucial manner by the volatility of the Brownian motion term in the wealth dynamics equation (9).

3. Simulations

In this section we present a simulation analysis of inter-generational wealth dynamics. We calibrate each of the models of Section 2 and compare their simulated wealth dynamics along various relevant empirical dimensions regarding the wealth distribution and wealth-rank persistence over generations. More precisely, we consider (*i*) the approximated rank-based model (1) calibrated using the standard model (5); and (*ii*) the permanently heterogeneous rank-based model (9), the extension of the rank-based model that includes permanent heterogeneity.

3.1. *Calibration*

In this section we discuss the details of the calibrations we adopt. We start from a parametrisation of the standard model mostly based on that in Benhabib *et al.* (2019). This parametrisation is the outcome of a simulated method of moments estimation procedure to match the wealth distribution and inter-generational social mobility data for the United States.[8] According to the estimates, we set the household lifespan T equal to 45 years and the growth rate of labour earnings equal to 0.01. The preference parameters η, ψ and χ are set equal to 0.04, 2 and 0.25, respectively. The estate tax and the capital income tax, b and ζ, are set equal to 0.2 and 0.15, respectively. To model yearly base labour income $y_{i,t}$, we use a six-state Markov chain calibrated using inter-generational persistence in labour income data from Chetty *et al.* (2014) together with data from the 2007 U.S. Survey of Consumer Finances (SCF) according to Díaz-Giménez *et al.* (2011).[9] Finally, for the idiosyncratic yearly return on wealth, $r_{i,t}$, we use a four-state Markov chain that is calibrated so that the average and standard deviation of these returns approximately match the empirical results of Fagereng *et al.* (2020) for Norwegian data.[10]

3.1.1. *The approximated rank-based model*

We then construct the rank-based approximation of the parametrisation of the standard model we just described. This approximation is obtained using (6) to define the rank-based parameters α_k and σ_k from (1). We first simulate the parametrisation of the standard model, which we do for one thousand generations with the number of households N set equal to 10,000. Importantly, this parametrisation induces by construction a stationary wealth distribution that matches the data for the United States well. We then use the results of these simulations and follow the econometric procedure described by Fernholz (2017) to estimate the expected relative growth-

[8] Wealth shares data are from the Survey of Consumer Finances, while inter-generational mobility data are from Charles and Hurst (2003).

[9] Earnings persistence by itself can hardly induce the wealth inequality observed in the data. This is because the distribution of earnings has a much thinner right tail than the distribution of wealth; see Benhabib *et al.* (2017); Benhabib and Bisin (2018). Thick wealth tails can be attained with stochastic returns across generations, as in this paper; or with stochastic discount factors, as in Krusell and Smith (1998) (indeed, it is the product or return and discount that matters in optimal accumulation decisions). Importantly, stochastic mortality with an exponential age distribution that has an independent constant probability of death or retirement ('perpetual youth' models) can also generate thick-tailed wealth distributions; see Castañeda *et al.* (2003), Benhabib and Bisin (2018, Section 3.4.1), Benhabib *et al.* (2016, Section 3.1) or Sargent *et al.* (2021). However, in this class of models wealth increases with age and the very rich tend to be unrealistically old (in the right tail of the age distribution): for example, an exponential age distribution with a calibrated constant death probability of 0.0167, implying an expected working life of 60 years, also implies that 15% of the working population have a working life of 114 years or more; see also Benhabib and Bisin (2018, pp. 1279–80).

[10] Specifically, we have $r_{i,t} \in \{0.02, 0.05, 0.09, 0.27\}$, with independent and identically distributed transition probabilities for the four states equal to (0.44, 0.45, 0.10, 0.01), respectively. With this parametrisation, the average and standard deviation of idiosyncratic returns are 4.3% and 3.1%, respectively. The estimates of the process for $r_{i,t}$ in Benhabib *et al.* (2019) are very close to these.

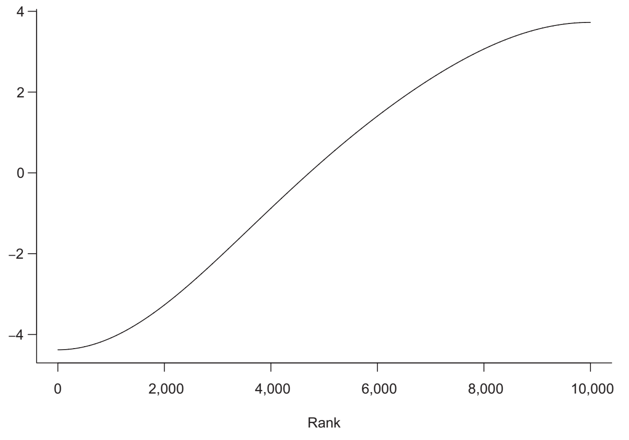

Fig. 1. *Annualised Estimated Parameters α_k for the Approximated Rank-Based Model.*

rate parameters α_k, $k = 1, \ldots, N$.[11] Using our estimates of the rank-based expected relative growth rate parameters α_k, we can find values for rank-based variance parameters σ_k satisfying (3) that, according to characterisation (4), yield a stationary distribution for the rank-based model that best approximates the average distribution of the standard model across the one thousand generations.[12]

Figure 1 plots the annualised estimated expected relative growth-rate parameters α_k for the rank-based approximation of the standard model. The figure shows that the these parameters satisfy the stability condition (2), with the estimated values such that $\alpha_1 < \alpha_2 < \cdots < \alpha_N$.[13]

Figure 2 plots the annualised estimated variance parameters σ_k for the rank-based approximation of the standard model. The estimated parameters α_k and σ_k from Figures 1 and 2 show a large difference in the growth rate of wealth for high- versus low-ranked households as well as relatively high wealth volatility for all households, especially for higher-ranked households. These are a consequence of the simultaneously high wealth inequality and low inter-generational correlation of wealth ranks in the data. The higher variance of wealth growth at higher ranks shown in Figure 2 is qualitatively consistent with the finding of Bach *et al.* (2020) that high-net-worth households in Sweden have more wealth volatility than other households.

Figure 3 presents a log-log plot of wealth versus rank for both the standard model and its rank-based approximation. This figure shows that the rank-based approximation generates a smoothed version of the wealth distribution from the standard model. Finally, Table 1 reports the wealth distribution shares in the data and those implied by the two models at their stationary distribution.

[11] Following Fernholz (2017), we apply a Gaussian kernel filter with a range of three thousand ranks ten times to smooth the estimated parameters α_k.

[12] Specifically, we minimise the squared distance between the wealth shares reported in Table 2 for the standard model and those predicted by (4) for the rank-based model.

[13] Recall from Subsection 3.1 and the paragraph just before Proposition 2, however, that the negative correlation between the growth rate of wealth and rank relation shown in Figure 1 does not necessarily imply that returns on wealth in the model are lower for higher-ranked, higher wealth households. In fact, returns on wealth are independent of rank in both the standard model and its rank-based approximation. Furthermore, recall that the α_k represent expected growth rates and hence their negative correlation with wealth ranks is consistent with higher wealth ranks being populated by individuals with high *realised* wealth returns and growth rates, which is indeed the case in the simulations.

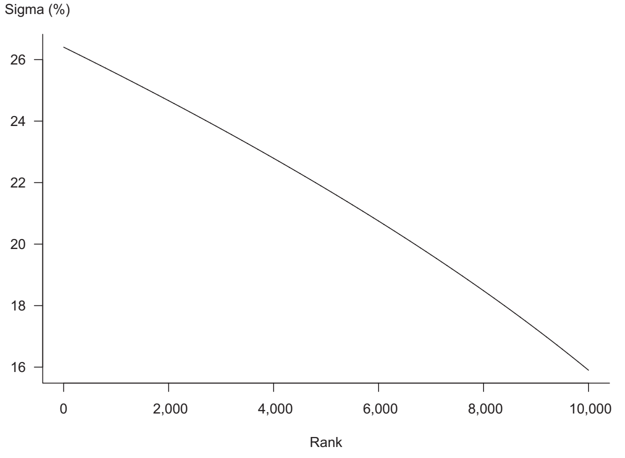

Fig. 2. *Annualised Estimated Parameters σ_k for the Approximated Rank-Based Model.*

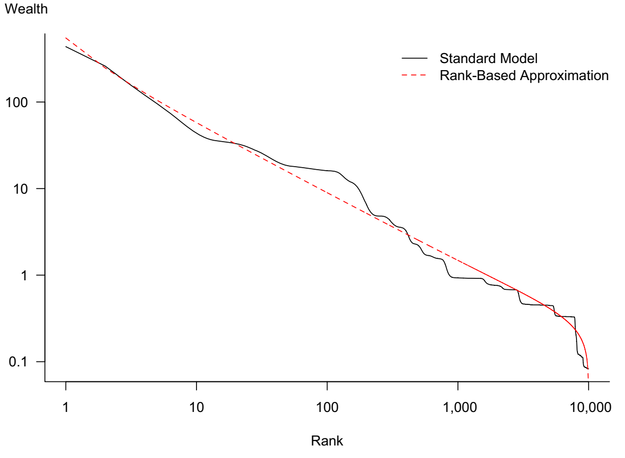

Fig. 3. *Log-Log Plot of Wealth versus Rank for the Standard Model (Average from One Thousand Simulations) and Its Rank-Based Approximation.*

Table 1. *Average Wealth Shares from One Thousand Simulations of the Different Models.*

	Data	Standard model	Approximated rank-based model
Wealth distribution			
Top 1%	33.6%	33.0%	31.9%
Top 1%–5%	26.7%	23.0%	17.1%
Top 5%–10%	11.1%	6.9%	9.5%
Top 10%–20%	12.0%	8.6%	11.2%
Top 20%–40%	11.2%	11.4%	13.1%
Top 40%–60%	4.5%	8.4%	8.3%
Bottom 40%	0.9%	8.6%	8.9%

Source. Data from the 2007 Survey of Consumer Finances.

The rank-based model closely approximates the standard model and both fit the data relatively well.

3.1.2. *The permanently heterogeneous rank-based model*

We calibrate the permanently heterogeneous rank-based model (9), which extends the rank-based model (1) to include permanent heterogeneity, so as to maintain approximately the same realistic stationary wealth distribution as the approximated rank-based model. We assume that three thousand of the households are high-growth households, with $\gamma_h = 0.015$. According to (10), this implies that the remaining seven thousand low-growth households have $\gamma_\ell \approx -0.0064$. Furthermore, we can use the same estimated parameter values for σ_k from the approximated rank-based model (Figure 2) for the permanently heterogeneous rank-based model.

Given the postulated distribution of γ_i and σ_k, the calibration of the rank-based relative growth rates $\hat{\alpha}_k$ is chosen to produce a stationary distribution similar to that produced by the standard and the approximated rank-based models (which, in turn, match well the distribution in the data). Indeed, we cannot simply use the estimated values of α_k from the approximated rank-based model (Figure 1) for the permanently heterogeneous rank-based model since the permanently heterogeneous parameters γ_i from (9) lead to a more skewed stationary distribution than in model (1).

Consider the rank-based approximation (1) of the heterogeneous rank-based model (9), where the parameters α_k are defined as in (6). In this case, Fernholz *et al.* (2013) showed that the relative growth rate parameters α'_k for the rank-based approximation are given by

$$\alpha'_k = \hat{\alpha}_k + (N - n)\xi_{\ell,k}\gamma_\ell + n\xi_{h,k}\gamma_h \tag{14}$$

for all $k = 1, \ldots, N$. According to Proposition 1, the stationary distributions of the rank-based approximation of model (9) and the rank-based model (1) will be the same if we choose $\hat{\alpha}_k$ such that $\alpha'_k = \alpha_k$ for each rank k. However, solving for the parameters $\hat{\alpha}_k$ that achieve this equality is complicated by the fact that we cannot directly solve for the occupation times $\xi_{\ell,k}$ and $\xi_{h,k}$ in (14), but instead must rely on simulations of the permanently heterogeneous rank-based model to generate estimates of these parameters.

We use a simple procedure to generate estimates of the parameters $\hat{\alpha}_k$ from model (9) such that α'_k is approximately equal α_k for each rank k. First, we use (14) to guess values of the parameters $\hat{\alpha}_k$ such that $\alpha'_k - \alpha_k \approx 0$ for all $k = 1, \ldots, N$. Next, we simulate the permanently heterogeneous rank-based model with these parameters $\hat{\alpha}_k$ to generate estimates of the rank-based approximation parameters α'_k, and then calculate the sum of squared errors of $\alpha'_k - \alpha_k$. Once this error term is calculated, we incrementally alter the values of $\hat{\alpha}_k$ by setting each equal to $x\hat{\alpha}_k$, where x is slightly less than or slightly greater than one. We then re-estimate the parameters α'_k and again calculate the sum of squared errors of $\alpha'_k - \alpha_k$. If the squared error with the parameter values $x\hat{\alpha}_k$ is smaller then we keep the new parameter values and repeat the procedure by altering the new parameter values in the same way. If not then we consider a different value of x and repeat the procedure. This procedure is repeated until the sum of squared errors of $\alpha'_k - \alpha_k$ is larger for the parameter values $x\hat{\alpha}_k$ for both $x = 1.001$ and $x = 0.999$. The annualised estimated parameters $\hat{\alpha}_k$ found using this procedure are shown in Figure 4.

3.2. *Results*

All the models we calibrate are stationary and hence can be simulated to generate stationary distributions of wealth that can be compared with the SCF data on wealth shares by percentile.

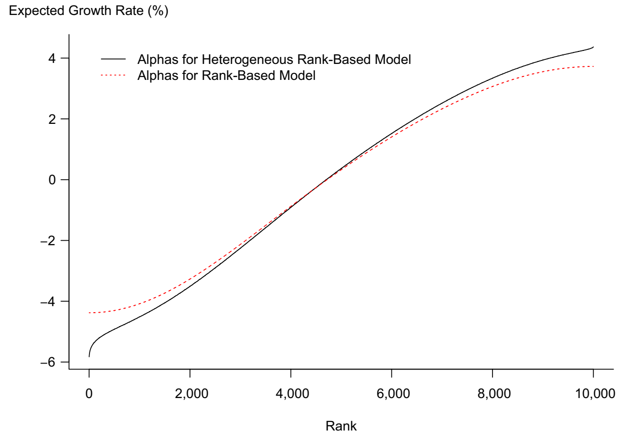

Fig. 4. *Annualised Estimated Parameters $\hat{\alpha}_k$ for the Permanently Heterogeneous Rank-Based Model, and Annualised Estimated Parameters α_k for the Approximated Rank-Based Model.*

Table 2. *Average Wealth Shares and Regression Coefficients form One Thousand Simulations of the Different Models.*

	Data	Approximated rank-based model	Perman. heterog. rank-based model
Wealth distribution			
Top 1%	33.6%	31.9%	34.0%
Top 1%–5%	26.7%	17.1%	16.6%
Top 5%–10%	11.1%	9.5%	9.2%
Top 10%–20%	12.0%	11.2%	10.8%
Top 20%–40%	11.2%	13.1%	12.7 %
Top 40%–60%	4.5%	8.3%	8.1%
Bottom 40%	0.9%	8.9%	8.5%
Wealth-rank correlations			
Parent-child rank coefficient	0.191	0.229	0.255
Grandparent-child rank coefficient	0.116	0.018	0.077
Long-run persistence coefficient	0.105	0.000	0.100

Notes: Reported coefficients are from regressions of child rank on parent rank and grandparent rank and from regressions of household rank in generation t on household rank in generation t - 23 (585 years). Data are from the 2007 Survey of Consumer Finances, Danish wealth holdings across three generations as reported by Boserup *et al.* (2014), and estimates of very long-run (585 years) dynastic wealth holdings in Florence, Italy, as reported by Barone and Mocetti (2016).

The results of these simulations are reported in the upper part of Table 2. Since these models are calibrated from a parametrisation of the standard model constructed to match these wealth shares, they all do relatively well at this, especially for the top 1% wealth share.

In addition to realistic wealth distributions, these models generate (*i*) parent-child and grandparent-child wealth-rank correlations (average coefficients from regressions of child rank on parent rank and grandparent rank—see the note to Table 2) that we compare to those of Boserup *et al.* (2014); and (*ii*) the long-run link of dynastic wealth ranks that we compare to those of Barone and Mocetti (2016). The results of these comparisons are reported in the lower part of

Table 3. *Average Composition of the Top 1%, Top 5%, Bottom 50% and Bottom 25% of Households from One Thousand Simulations of the Heterogeneous Rank-Based Model.*

	Top 1%	Top 5%	Bottom 50%	Bottom 25%
High-growth households	82.6%	68.8%	18.3%	12.8%
Low-growth households	17.4%	31.2%	81.7%	87.2%

Table 2.[14] Both calibrated models tend to generate parent-child wealth-rank correlations slightly higher than in the data. They also tend to generate grandparent-child wealth-rank correlations lower than in the data— though the permanently heterogeneous rank-based model fares much better in this respect. With regards to long-run correlations, the results we obtain are consistent with the theoretical results of Section 2. As implied by Proposition 3, household wealth ranks are uncorrelated over very long time periods in the rank-based approximation of the standard model. Household wealth ranks are instead positively correlated over arbitrarily long time periods in a rank-based model that features permanent heterogeneity, as allowed by Theorem 1.[15] In conclusion, the permanent heterogeneity in the rank-based model helps to match rather well all aspects of the data simultaneously—the wealth distribution, the link between child, parent and grandparent wealth ranks, and the positive correlation of dynastic wealth ranks over very long time periods.

It is useful, then, to study the properties of the heterogeneous rank-based model more closely. Table 3 shows the composition of the top 1% and top 5% wealth-ranked households in terms of low- and high-growth households. This table also shows the composition of the bottom 50% and bottom 25% ranked households. According to the table, high-growth households make up the great majority of the top 1% and 5%, but there is still a non-negligible minority of low-growth households in these top subsets. The results in the table also suggest that low-growth households are more common in top subsets of the wealth distribution than high-growth households are in bottom subsets of the wealth distribution. Indeed, the fraction of low-growth households in the bottom 25% approximately matches the fraction of high-growth households in the top 1%, even though the latter is a much smaller and more exclusive subset of the wealth distribution.

Figures 5 and 6 plot the estimated occupation times of different percentiles of the wealth distribution form high- and low-growth households, respectively. The estimated occupation times presented in the figures are clearly consistent with the result in Proposition 4. Because there are three thousand high-growth households and seven thousand low-growth households, the maximum average occupation time for a high-growth household in any percentile of the wealth distribution is $1/3{,}000 \approx 0.033\%$, while the maximum occupation time for a low-growth household in any percentile is $1/7{,}000 \approx 0.014\%$.[16]

[14] The data we use to evaluate the models refer to different countries, though all are developed market economies. This is due to a lack of comparable evidence for the United States. It is arguably not problematic in that we simply aim at a general theoretical and empirical understanding of the fundamental elements of a model of wealth dynamics rather than at a formal estimation procedure. Interestingly, estimates of returns to wealth in the United States (Benhabib *et al.*, 2019) and in Norway (Fagereng *et al.*, 2020) are very close; and several of these developed market economies in the West tend to share comparable wealth distributions, at least in terms of their inequality (measured by the Gini coefficient); see Benhabib *et al.* (2017).

[15] The standard model is, like its rank-based approximation, unable to generate long-run wealth-rank correlations as well.

[16] These upper bounds for low- and high-growth household occupation times also appear in Proposition 4, since the number of high-growth households in this simulation n is equal to 3,000.

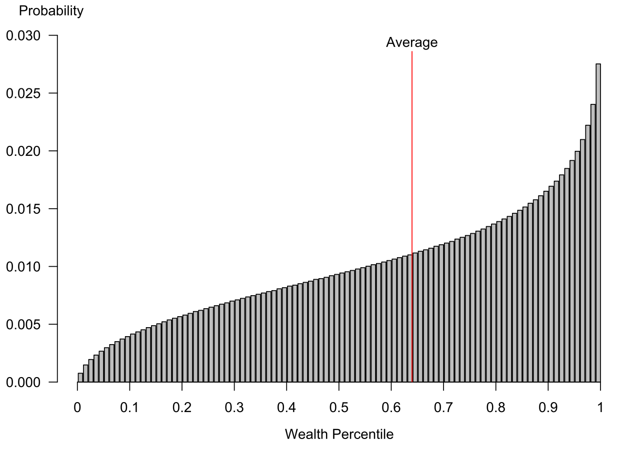

Fig. 5. *Average High-Growth Household Occupation Times for Different Percentiles of the Wealth Distribution from One Thousand Simulations of the Heterogeneous Rank-Based Model.*

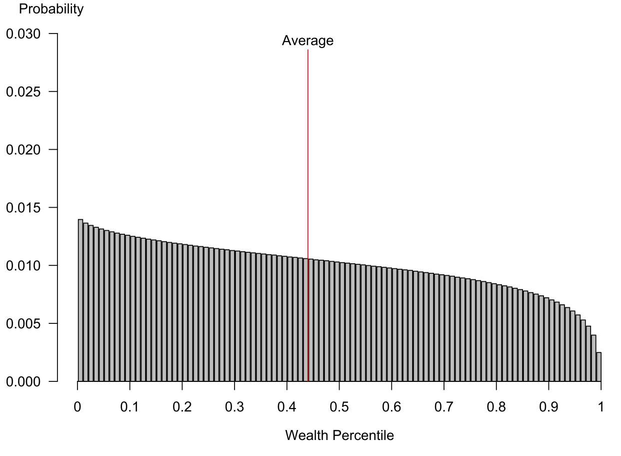

Fig. 6. *Average Low-Growth Household Occupation Times for Different Percentiles of the Wealth Distribution from One Thousand Simulations of the Heterogeneous Rank-Based Model.*

3.2.1. *Sensitivity analysis: auto-correlated returns*

To better identify the role of permanent heterogeneity in capturing both the wealth distribution and wealth-rank persistence over generations, in this section we compare it with a different form of imperfect social mobility, inter-generationally auto-correlated returns to wealth.[17] More specifically we report on the simulations of an extension of the standard model (5) in which there is no permanent component to the growth rate of wealth but returns to wealth are highly auto-correlated across generations.

[17] The persistent heterogeneity in household-saving behaviour is the mechanism exploited by Degan and Thibault (2016) to induce long correlations across generations. Such a mechanism is also at work in both the permanently heterogeneous and the auto-correlated returns models, as dynasties with higher returns endogenously display a higher savings rate.

Table 4. *Average Wealth Shares and Regression Coefficients form One Thousand Simulations of the Different Models.*

	Data	Auto-correlated returns model ($\theta = 0.95$)	Perman. heterog. rank-based model
Wealth distribution			
Top 1%	33.6%	31.5%	34.0%
Top 1%–5%	26.7%	20.6%	16.6%
Top 5%–10%	11.1%	12.3%	9.2%
Top 10%–20%	12.0%	13.5%	10.8%
Top 20%–40%	11.2%	12.8%	12.7%
Top 40%–60%	4.5%	5.8%	8.1%
Bottom 40%	0.9%	3.5%	8.5%
Wealth-rank correlations			
Parent-child rank coefficient	0.191	0.407	0.255
Grandparent-child rank coefficient	0.116	0.044	0.077
Long-run persistence coefficient	0.105	0.041	0.100

Notes: See the Notes in Table 2.

We assume that wealth returns follow a highly persistent AR-1 process, with

$$\log(1 + r_{i,t+1}) = \theta \log(1 + r_{i,t}) + \epsilon_{i,t},$$

where $\epsilon_{i,t}$ is normally distributed with mean equal to 0.041 and the persistence parameter θ is equal to 0.95.[18] The standard deviation of $\epsilon_{i,t}$ is chosen to match the U.S. wealth distribution data according to the 2007 SCF (Díaz-Giménez *et al.*, 2011). For symmetry, we assume that labour earnings $\log y_{i,t}$ also follow an AR-1 process with persistence equal to 0.3, which matches the inter-generational persistence of income in the United States according to Chetty *et al.* (2014), and with mean and standard deviation chosen to match the calibration of income in the standard model.[19] In all other respects, the auto-correlated returns model is identical to the standard model that was used to calibrate the rank-based model.

We report the results of these simulations in Table 4. These results show that the version of the standard model with highly auto-correlated returns is able to match rather well the wealth shares in the data. Interestingly, it is able to generate a significant grandparent-child wealth-rank correlation, but to do so, it requires a much too strong correlation between the parent and the child wealth ranks. Fundamentally, however, the auto-correlated returns model fails to match the long-run persistence coefficient in wealth ranks that instead is quite precisely captured by the permanently heterogeneous rank-based model.

To better understand these results, the long-run persistence of wealth ranks implied by the permanently heterogeneous rank-based model and the auto-correlated returns model are compared most clearly in Figure 7. In this figure, we plot the correlation between the wealth ranks of households in generation t and generation $t + x$, with values of x ranging from 1 to 25, for both the heterogeneous rank-based and auto-correlated returns models. Although the auto-correlated returns model is able to generate substantial persistence in rank across one or two generations, the rank correlation in this model quickly declines towards zero as the generational gap between

[18] Results are very similar for $\theta = 0.90$.

[19] Note that the standard model is calibrated so that income matches the U.S. data according to the 2007 SCF.

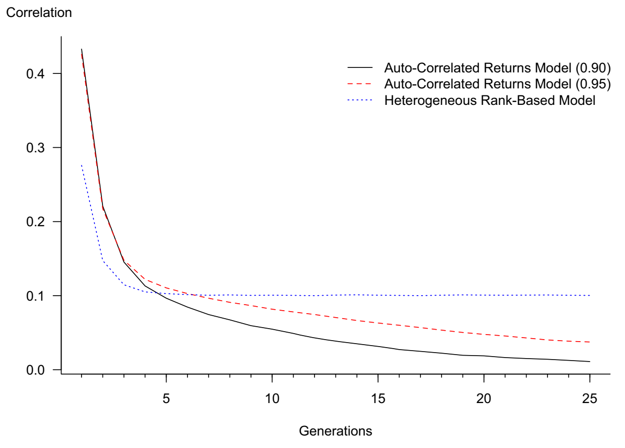

Fig. 7. *Rank Correlations Across Multiple Generations from One Thousand Simulations of the Permanently Heterogeneous Rank-Based and Auto-Correlated Returns Models.*

households increases. In contrast, the permanently heterogeneous rank-based model generates a more realistic but smaller persistence in wealth rank across one or two generations, and this persistence never falls below 0.1 even as the generational gap grows large. Of course, this very long-run persistence in wealth rank is exactly what is predicted by Theorem 1.

Another important dimension along which it is useful to compare the predictions of the permanently heterogeneous rank-based model and the auto-correlated returns models is parent-child return-rank coefficients. These correlation coefficients are interesting also as possible indicators of the mechanisms behind long-run wealth persistence. A relatively high parent-child return-rank coefficient suggests direct inter-generational mechanisms, like cultural transmission. A low coefficient suggests, on the contrary, institutional factors, like the perpetuation of the political and economic elites, which are extremely persistent but do not run directly from parent to child (Bourdieu, 1984; 1998; Acemoglu and Robinson, 2008; Bisin and Verdier, 2017). Using Norwegian data, Fagereng *et al.* (2020) reported a small coefficient, 0.16, from a regression of child return rank on parent return rank.

Both the permanently heterogeneous and the auto-correlated returns models induce by construction some parent-child correlations in the returns to wealth.[20] In both models however this correlation is reduced by the postulated volatility of the growth rate, which is captured by the term σ_k in (9) for the permanently heterogeneous model and by the variance of $\epsilon_{i,t}$ in the auto-correlated model. In both models, these variances are chosen to match the distribution of wealth and wealth-rank correlations as reported in Table 4. Interestingly, we find that the permanently

[20] In order to calculate the parent-child return-rank coefficient for the permanently heterogeneous rank-based model, it is necessary to decompose the parameters $\hat{\alpha}_k$ from (9) into log return and log labour income to capital income components, as in (8). Specifically, we can write (9) as

$$d \log w_i(t) = (\gamma_i + \kappa_{\rho_t(i)} + \omega_{\rho_t(i)})\, dt + \sigma_{\rho_t(i)}\, dB_i(t),$$

where the parameters κ_k and ω_k respectively measure the log return to wealth and the log ratio of labour income to capital income for the kth ranked household and satisfy $\hat{\alpha}_k = \kappa_k + \omega_k$ for each rank $k = 1, \dots, N$. We assume that returns to wealth are constant across wealth ranks as in the standard model. As a consequence, $\kappa_k = 0$ for all $k = 1, \dots, N$ (following the normalisation of the $\hat{\alpha}_k$ parameters, and without loss of generality, we normalise the κ_k parameters to sum to zero).

heterogeneous rank-based model, in our calibration, produces quite a small parent-child correlation of returns, even smaller than in the Norwegian data, while this correlation is much higher for the auto-correlated returns model. Specifically, the coefficient from a regression of child return rank on parent return rank, averaged across one thousand simulations, is equal to 0.08 for the permanently heterogeneous model[21] and is equal to 0.95 for the auto-correlated returns model. Effectively, in our calibration, the volatility of the growth rate in the permanently heterogeneous model introduces enough churning in parent-child returns to lower their rank correlation substantially. This is not the case for the auto-correlated model, which requires a small volatility of the growth rate to match the distribution of wealth and wealth-rank correlations and, as a consequence, dramatically misses the low return-rank correlation documented by Fagereng *et al.* (2020).

4. Conclusion

We consider a simple heterogeneous agents model based on Benhabib *et al.* (2019) and show that such standard models fail to match recent empirical results regarding long-run wealth mobility. In particular, this type of model does not generate a positive correlation between grandparent-child wealth rank, after controlling for parent-child wealth rank, and does not generate a positive correlation between dynastic wealth ranks across very long time periods. We extend the standard model to include permanent heterogeneity in returns to wealth, and show that such an extended model is able to simultaneously match the wealth distribution, short-run wealth mobility and long-run wealth mobility. Finally, we find that the model with permanent heterogeneity produces in our calibration a small parent-child correlation of returns, close to that documented by Fagereng *et al.* (2020) for Norway. This suggests persistent institutional factors as mechanisms for sustaining long-run wealth persistence, rather than direct inter-generational mechanisms like cultural transmission. While we do not have enough structure and data to identify particular institutional channels responsible for the long-run persistence in wealth correlations, future work along these lines is needed to further this literature.

New York University, USA
New York University, USA
Claremont McKenna College, USA

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Acemoglu, D. and Robinson, J.A. (2008). ‘Persistence of power, elites, and institutions’, *American Economic Review*, vol. 98(1), pp. 267–93.

Bach, L., Calvet, L.E. and Sodini, P. (2020). ‘Rich pickings? Risk, return, and skill in household wealth’, *American Economic Review*, vol. 110(9), pp. 2703–47.

Banner, A., Fernholz, R. and Karatzas, I. (2005). ‘Atlas models of equity markets’, *Annals of Applied Probability*, vol. 15(4), pp. 2296–330.

[21] We conjecture that, using a model that incorporates higher returns at higher wealth ranks as in Benhabib *et al.* (2019) to estimate a different permanently heterogeneous rank-based model of the form (9), could generate an even closer match to the 0.16 parent-child return-rank coefficient in Fagereng *et al.* (2020).

Barone, G. and Mocetti, S. (2016). 'Intergenerational mobility in the very long run: Florence 1427–2011', Working Paper 1060, Bank of Italy.

Benhabib, J. and Bisin, A. (2018). 'Skewed wealth distributions: Theory and empirics', *Journal of Economic Literature*, vol. 56(4), pp. 1261–91.

Benhabib, J., Bisin, A. and Luo, M. (2017). 'Eearnings inequality and other determinants of wealth inequality', *American Economic Review: Papers and Proceedings*, vol. 107(5), pp. 593–7.

Benhabib, J., Bisin, A. and Luo, M. (2019). 'Wealth distribution and social mobility in the US: A quantitative approach', *American Economic Review*, vol. 109(5), pp. 1623–47.

Benhabib, J., Bisin, A. and Zhu, S. (2011). 'The distribution of wealth and fiscal policy in economies with finitely lived agents', *Econometrica*, vol. 79(1), pp. 123–57.

Benhabib, J., Bisin, A. and Zhu, S. (2016). 'The distribution of wealth in the Blanchard-Yaari model', *Macroeconomic Dynamics*, vol. 20(2), pp. 466–81.

Bisin, A. and Moro, A. (2021). 'Late for history', in (A. Bisin and G. Federico, eds.), *Handbook of Historical Economics*, pp. 269–96, London: Academic Press.

Bisin, A. and Verdier, T. (2001). 'The economics of cultural transmission and the dynamics of preferences', *Journal of Economic Theory*, vol. 97(2), pp. 298–319.

Bisin, A. and Verdier, T. (2017). 'On the joint evolution of culture and institutions', Technical Report, National Bureau of Economic Research.

Bisin, A. and Verdier, T. (2021). 'Phase diagrams in historical economics: Culture and institutions', in (A. Bisin and G. Federico, eds.), *Handbook of Historical Economics*, Working Paper 23375, London: Academic Press.

Boserup, S., Kopczuk, W. and Kreiner, C. (2014). 'Stability and persistence of intergenerational wealth formation: Evidence from Danish wealth records of three generations', Working Paper, University of Copenhagen.

Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgement of Taste*, Cambridge, MA: Harvard University Press.

Bourdieu, P. (1998). *The State Nobility: Elite Schools in the Field of Power*, Stanford, CA: Stanford University Press.

Bourdieu, P. and Passeron, J. (1970). *La reproduction: éléments pour une théorie du système d'enseignement*, Paris: Editions de Minuit.

Braun, S.T. and Stuhler, J. (2018). 'The transmission of inequality across multiple generations: Testing recent theories with evidence from Germany', ECONOMIC JOURNAL, vol. 128(609), pp. 576–611.

Cagetti, M. and De Nardi, M. (2006). 'Entrepreneurship, frictions, and wealth', *Journal of Political Economy*, vol. 114(5), pp. 835–70.

Castañeda, A., Díaz-Giménez, J. and Ríos-Rull, J.-V. (2003). 'Accounting for U.S. earnings and wealth inequality', *Journal of Political Economy*, vol. 111(4), pp. 818–57.

Chan, T.W. and Boliver, V. (2013). 'The grandparents effect in social mobility: Evidence from British birth cohort studies', *American Sociological Review*, vol. 78(4), pp. 662–78.

Charles, K.K. and Hurst, E. (2003). 'The correlation of wealth across generations', *Journal of Political Economy*, vol. 111(6), pp. 1155–82.

Chetty, R., Hendren, N., Kline, P. and Saez, E. (2014). 'Where is the land of opportunity? The geography of intergenerational mobility in the United States', *Quarterly Journal of Economics*, vol. 129(4), pp. 1553–623.

Clark, G. (2014). *The Son Also Rises: Surnames and the History of Social Mobility*, Princeton, NJ: Princeton University Press.

Clark, G. and Cummins, N. (2015). 'Intergenerational wealth mobility in England, 1858–2012: Surnames and social mobility', ECONOMIC JOURNAL, vol. 125(582), pp. 61–85.

Degan, A. and Thibault, E. (2016). 'Dynastic accumulation of wealth', *Mathematical Social Sciences*, vol. 81(C), pp. 66–78.

Díaz-Giménez, J., Glover, A. and Ríos-Rull, J.-V. (2011). 'Facts on the U.S. distributions of earnings, income, and wealth in the United States: 2007 update', *Federal Reserve Bank of Minneapolis Quarterly Review*, vol. 34(1), pp. 2–31.

Fagereng, A., Guiso, L., Malacrino, D. and Pistaferri, L. (2020). 'Heterogeneity and persistence in returns to wealth', *Econometrica*, vol. 88(1), pp. 115–70.

Fernholz, R.T. (2017). 'Nonparametric methods and local-time-based estimation for dynamic power law distributions', *Journal of Applied Econometrics*, vol. 32(7), pp. 1244–60.

Fernholz, R., Ichiba, T. and Karatzas, I. (2013). 'A second-order stock market model', *Annals of Finance*, vol. 9(3), pp. 439–54.

Hertel, F.R. and Groh-Samberg, O. (2014). 'Class mobility across three generations in the U.S. and Germany', *Research in Social Stratification and Mobility*, vol. 35, pp. 35–52.

Hubmer, J., Krusell, P. and Smith, A.A., Jr. (2016). 'The historical evolution of the wealth distribution: A quantitative-theoretic investigation', Working Paper 23011, National Bureau of Economic Research.

Ichiba, T., Papathanakos, V., Banner, A., Karatzas, I. and Fernholz, R. (2011). 'Hybrid atlas models', *Annals of Applied Probability*, vol. 21(2), pp. 609–44.

Kindermann, F. and Krueger, D. (2021). 'High marginal tax rates on the top 1%? Lessons from a life cycle model with idiosyncratic income risk', *American Economic Journal: Macroeconomics*. Available at: https://www.aeaweb.org/articles?id=10.1257/mac.20150170&&from=f as AMERICAN ECONOMIC JOURNAL: MACROECONOMICS (FORTHCOMING).

Krusell, P. and Smith, A.A. (1998). ‘Income and wealth heterogeneity in the macroeconomy’, *Journal of Political Economy*, vol. 106(5), pp. 867–96.
Krusell, P. and Smith, A.A. (2015). ‘Is Piketty’s ‘second law of capitalism’ fundamental?’, *Journal of Political Economy*, vol. 123(4), pp. 725–48.
Lindahl, M., Palme, M., Massih, S.S. and Sjögren, A. (2015). ‘Long-term intergenerational persistence of human capital: An empirical analysis of four generations’, *Journal of Human Resources*, vol. 50(1), pp. 1–33.
Long, J. and Ferrie, J. (2013). ‘Intergenerational occupational mobility in Great Britain and the United States since 1850’, *American Economic Review*, vol. 103(4), pp. 1109–37.
Mare, R. (2011). ‘A multigenerational view of inequality’, *Demography*, vol. 48, pp. 1–23.
Modin, B., Erikson, R. and Vagero, D. (2013). ‘Intergenerational continuity in school performance: Do grandparents matter?’, *European Sociological Review*, vol. 29(4), pp. 858–70.
Quadrini, V. (2000). ‘Entrepreneurship, saving, and social mobility’, *Review of Economic Dynamics*, vol. 3(1), pp. 1–40.
Saporta, B. and Yao, J.-F. (2005). ‘Tail of a linear diffusion with Markov switching’, *Annals of Applied Probability*, vol. 15(1B), pp. 992–1018.
Sargent, T.J., Wang, N. and Yang, J. (2021). ‘Earnings growth and the wealth distribution’, *Proceedings of the National Academy of Sciences*, vol. 118(15).
Stuhler, J. (2012). ‘Mobility across multiple generations: The iterated regression fallacy’, Discussion Paper 7072, Institute of Labor Economics.
Voth, H.-J. (2021). ‘Persistence—myth and mystery’, in (A. Bisin and G. Federico, eds.), *Handbook of Historical Economics*, pp. 243–67, London: Academic Press.
Warren, J.R. and Hauser, R.M. (1997). ‘Social stratification across three generations: New evidence from the Wisconsin longitudinal study’, *American Sociological Review*, vol. 62(4), pp. 561–72.