Online Appendix

APPENDIX C: PROOFS FOR APPENDIX C

PROOF OF PROPOSITION 6: Denote $\boldsymbol{\mu}_t(w) = (\mu_{jt}(w))_{1\leq j\leq J}$ the vector of drifts for each type. Similarly, denote $\boldsymbol{\nu}_t(w) = (\nu_{jt}(w))_{1\leq j\leq J}$ the vector of idiosyncratic volatilities for each type. The Kolmogorov forward equation for $\mathbf{g}_t$ is

$$\mathrm{d}\mathbf{g_t}(w) = -\partial_w(\boldsymbol{\mu}_t(w)w\cdot\mathbf{g}_t(w))\,\mathrm{d}t + \frac{1}{2}\partial_w^2\left(\boldsymbol{\nu}_t(w)w\cdot\boldsymbol{\nu}_t(w)w\cdot\mathbf{g}_t(w)\right)\mathrm{d}t + \mathbb{T}'\mathbf{g}_t(w)\,\mathrm{d}t,$$

where $\cdot$ denotes the element-wise product and $\mathbb{T}$ denotes the Markov transition matrix across the groups.

Denote $g_t = \sum_{j=1}^J g_{jt}$ the overall wealth density at time t. Using the previous equation, the law of motion of the overall wealth density is:

$$\mathrm{d}g_t(w) = -\partial_w\left((\boldsymbol{\mu}_t(w)w)'\mathbf{g}_t(w)\right)\mathrm{d}t + \frac{1}{2}\partial_w^2\left((\boldsymbol{\nu}_t(w)w\cdot\boldsymbol{\nu}_t(w)w)'\mathbf{g}_t(w)\right)\mathrm{d}t.$$

As in the proof of Proposition 2, we can plug this equation into (22) and integrate by parts to obtain

$$\mathrm{d}\left(\int_{q_t}^\infty wg_t(w)\,\mathrm{d}w\right) = \left(\int_{q_t}^\infty(\boldsymbol{\mu}_t(w)w)'\mathbf{g}_t(w)\,\mathrm{d}w\right)\mathrm{d}t + \frac{1}{2}\left(\boldsymbol{\nu}_t(q_t)q_t\cdot\boldsymbol{\nu}_t(q_t)q_t\right)'\mathbf{g}_t(q_t)\,\mathrm{d}t.$$

Dividing by $\int_{q_t}^\infty wg_t(w)\,\mathrm{d}w$ gives the result. *Q.E.D.*

PROOF OF PROPOSITION 7: I follow the same steps as in the proof of Proposition 2, with the key difference that the wealth density is now stochastic. As discussed in the text, the proof is only informal.

I start by rederiving the law of motion of the quantile q_t in the presence of aggregate risk. Applying Ito's lemma on the implicit definition of the quantile (20) gives

$$0 = \int_{q_t}^\infty \mathrm{d}g_t(w)\,\mathrm{d}w - g_t(q_t)\,\mathrm{d}q_t - \sigma[\mathrm{d}g_t(q_t)]^2\,\mathrm{d}t,$$

where $\sigma[\mathrm{d}g_t(q_t)]$ denote the exposure of $g_t(q_t)$ to aggregate shocks. Applying Ito's lemma gives the law of motion of the total wealth in the top percentile:

$$\mathrm{d}\left(\int_{q_t}^\infty wg_t(w)\,\mathrm{d}w\right) = \int_{q_t}^\infty w\,\mathrm{d}g_t(w)\,\mathrm{d}w - q_tg_t(q_t)\,\mathrm{d}q_t - q_t\sigma[\mathrm{d}g_t(q_t)]^2\,\mathrm{d}t - \frac{1}{2}g_t(q_t)\sigma[\mathrm{d}q_t]^2\,\mathrm{d}t.$$

Combining these two equations gives

$$\begin{aligned}\mathrm{d}\left(\int_{q_t}^\infty wg_t(w)\,\mathrm{d}w\right) &= \int_{q_t}^\infty(w-q_t)\,\mathrm{d}g_t(w)\,\mathrm{d}w - \frac{1}{2}g_t(q_t)\sigma[\mathrm{d}q_t]^2\,\mathrm{d}t\\ &= \int_{q_t}^\infty(w-q_t)\,\mathrm{d}g_t(w)\,\mathrm{d}w - \frac{1}{2}\frac{1}{g_t(q_t)}\left(\int_{q_t}^\infty\sigma[\mathrm{d}g_t(w)]\,\mathrm{d}w\right)^2\mathrm{d}t. \quad (34)\end{aligned}$$

Denote $\boldsymbol{\sigma}_t(w) = (\sigma_{jt}(w))_{1\le j\le J}$ the vector giving the aggregate exposures of each type. As shown in Kurtz and Xiong (1999) and Carmona and Delarue (2018), in the presence of aggregate risk, the wealth density evolves as

$$\begin{aligned}\mathrm{d}\mathbf{g_t}(w) &= -\partial_w(\boldsymbol{\mu}_t(w)w\cdot\mathbf{g}_t(w))\,\mathrm{d}t - \partial_w(\boldsymbol{\sigma}_t(w)w\cdot\mathbf{g}_t(w))\,\mathrm{d}Z_t\\ &\quad+\frac{1}{2}\partial_w^2\left((\boldsymbol{\nu}_t(w)w\cdot\boldsymbol{\nu}_t(w)w+\boldsymbol{\sigma}_t(w)w\cdot\boldsymbol{\sigma}_t(w)w)\cdot\mathbf{g}_t(w)\right)\mathrm{d}t+\mathbb{T}'\mathbf{g}_t(w)\,\mathrm{d}t.\end{aligned}$$

Therefore, the law of motion of the overall wealth density, g_t, is

$$\begin{aligned}\mathrm{d}g_t(w) &= -\partial_w\left((\boldsymbol{\mu}_t(w)w)'\mathbf{g}_t(w)\right)\mathrm{d}t - \partial_w\left((\boldsymbol{\sigma}_t(w)w)'\mathbf{g}_t(w)\right)\mathrm{d}Z_t\\ &\quad+\frac{1}{2}\partial_w^2\left((\boldsymbol{\nu}_t(w)w\cdot\boldsymbol{\nu}_t(w)w+\boldsymbol{\sigma}_t(w)w\cdot\boldsymbol{\sigma}_t(w)w)'\mathbf{g}_t(w)\right)\mathrm{d}t.\end{aligned}$$

Plugging this equation into (34) and integrating by parts gives

$$\begin{aligned}\mathrm{d}\left(\int_{q_t}^{\infty} wg_t(w)\,\mathrm{d}w\right) &= \left(\int_{q_t}^{\infty}(\boldsymbol{\mu}_t(w)w)'\mathbf{g}_t(w)\,\mathrm{d}w\right)\mathrm{d}t + \left(\int_{q_t}^{\infty}(\boldsymbol{\sigma}_t(w)w)'\mathbf{g}_t(w)\,\mathrm{d}w\right)\mathrm{d}Z_t\\ &\quad+\frac{1}{2}q_t^2\left(\boldsymbol{\nu}_t(q_t)\cdot\boldsymbol{\nu}_t(q_t)+\boldsymbol{\sigma}_t(q_t)\cdot\boldsymbol{\sigma}_t(q_t)\right)'\mathbf{g}_t(q_t)\,\mathrm{d}t\\ &\quad-\frac{1}{2}\frac{1}{g_t(q_t)}q_t^2\left(\boldsymbol{\sigma}_t(q_t)'\mathbf{g}_t(q_t)\right)^2\mathrm{d}t.\end{aligned}$$

Dividing by $\int_{q_t}^{\infty} wg_t(w)\,\mathrm{d}w$ gives the result. *Q.E.D.*

PROOF OF PROPOSITION 8: I follow the same steps as in the proof of Proposition 2. The Kolmogorov forward equation associated with the process (29) is:

$$\begin{aligned}\mathrm{d}g_t(w) &= -\partial_w(\mu_t(w)wg_t(w))\,\mathrm{d}t+\frac{1}{2}\partial_w^2\left(\nu_t^2(w)w^2g_t(w)\right)\mathrm{d}t\\ &\quad+\Bigg(\int_0^{\infty}\frac{1}{\phi_t(w')w}f_U\left(\frac{\log(w/w')}{\phi_t(w')}\right)g_t(w')\,\mathrm{d}w'-g_t(w)\\ &\qquad+\partial_w\left(\mathrm{E}_U\left[e^{\phi_t(w)U}-1\right]wg_t(w)\right)\Bigg)\lambda\,\mathrm{d}t.\end{aligned}$$

Plugging this into (22) and integrating by parts gives

$$\begin{aligned}\mathrm{d}\left(\int_{q_t}^{\infty} wg_t(w)\,\mathrm{d}w\right) &= \left(\int_{q_t}^{\infty}\mu(w)wg_t(w)\,\mathrm{d}w\right)\mathrm{d}t+\frac{1}{2}g_t(q_t)q_t^2\nu_t(q_t)^2\,\mathrm{d}t\\ &\quad+\Bigg(\int_{q_t}^{\infty}(w-q_t)\left(\int_0^{\infty}\frac{1}{\phi_t(w')w}f_U\left(\frac{\log(w/w')}{\phi_t(w')}\right)g_t(w')\,\mathrm{d}w'-g_t(w)\right)\mathrm{d}w\\ &\qquad-\int_{q_t}^{\infty}\mathrm{E}_U\left[e^{\phi_t(w)U}-1\right]wg_t(w)\,\mathrm{d}w\Bigg)\lambda\,\mathrm{d}t.\end{aligned}$$

The last "jump" term can be rewritten as follows:

$$\mathrm{E}_U\left[\int_0^{\infty}\left(e^{\phi_t(w')U}w'-q_t\right)^+g_t(w')\,\mathrm{d}w'-\int_{q_t}^{\infty}\left(e^{\phi(w)U}w-q_t\right)g_t(w)\,\mathrm{d}w\right]\lambda\,\mathrm{d}t$$

$$= \mathrm{E}_U\left[\int_0^{q_t}\left(e^{\phi_t(w)U}w-q_t\right)^+ g_t(w)\,\mathrm{d}w+\int_{q_t}^{\infty}\left(q_t-e^{\phi_t(w)U}w\right)^+ g_t(w)\,\mathrm{d}w\right]\lambda\,\mathrm{d}t.$$

Dividing by $\int_{q_t}^{\infty} wg_t(w)\,\mathrm{d}w$ gives the result. *Q.E.D.*

PROOF OF COROLLARY 9: Using the change of variable $J=\phi_t(w)U$, one can rewrite the between term due to jump in (30) as

$$\begin{aligned}
&\frac{1}{pw_{\mathcal{P}t}}\left(\int_J\left(\int_0^{q_t}(e^Jw-q_t)^+\frac{1}{\phi_t(w)}f_U\left(\frac{J}{\phi_t(w)}\right)g(w)\,\mathrm{d}w\right.\right.\\
&\qquad\left.\left.+\int_{q_t}^{\infty}(q_t-e^Jw)^+\frac{1}{\phi_t(w)}f_U\left(\frac{J}{\phi_t(w)}\right)g(w)\,\mathrm{d}w\right)\mathrm{d}J\right)\lambda\,\mathrm{d}t\\
&=\frac{1}{pw_{\mathcal{P}t}}\left(\int_J\int_{q_te^{-J}}^{q_t}(e^Jw-q_t)\frac{1}{\phi_t(w)}f_U\left(\frac{J}{\phi_t(w)}\right)g_t(w)\,\mathrm{d}w\,\mathrm{d}J\right)\lambda\,\mathrm{d}t
\end{aligned}$$

Assuming that the function $v\to\int_{q_te^{-v}}^{q_t}(e^vw-q_t)\frac{1}{\phi_t(w)}f_U(\frac{J}{\phi_t(w)})g_t(w)\,\mathrm{d}w$ is analytic, we can rewrite the expression as follows:

$$\begin{aligned}
&\frac{1}{pw_{\mathcal{P}t}}\left(\int_J\sum_{n=2}^{\infty}\frac{1}{n!}\partial_v^n\left(\int_{q_te^{-v}}^{q_t}(e^vw-q_t)\frac{1}{\phi_t(w)}f_U\left(\frac{J}{\phi_t(w)}\right)g_t(w)\,\mathrm{d}w\right)\Big|_{v=0}J^n\,\mathrm{d}J\right)\lambda\,\mathrm{d}t\\
&=\frac{1}{pw_{\mathcal{P}t}}\left(\sum_{n=2}^{\infty}\frac{1}{n!}\partial_v^n\left(\int_{q_te^{-v}}^{q_t}(e^vw-q_t)\left(\int_J J^n\frac{1}{\phi_t(w)}f_U\left(\frac{J}{\phi_t(w)}\right)\mathrm{d}J\right)g_t(w)\,\mathrm{d}w\right)\Big|_{v=0}\right)\lambda\,\mathrm{d}t\\
&=\frac{1}{pw_{\mathcal{P}t}}\sum_{n=2}^{\infty}\frac{1}{n!}\partial_v^n\left(\int_{q_te^{-v}}^{q_t}(e^vw-q_t)\left(\int_U\lambda\phi_t(w)^nU^nf_U(U)\,\mathrm{d}U\right)g_t(w)\,\mathrm{d}w\right)\Big|_{v=0}\mathrm{d}t\\
&=\frac{1}{pw_{\mathcal{P}t}}\sum_{n=2}^{\infty}\frac{1}{n!}\partial_v^n\left(\int_{q_te^{-v}}^{q_t}(e^vw-q_t)\left(\lambda\phi_t(w)^n\mathrm{E}_U[U^n]\right)g_t(w)\,\mathrm{d}w\right)\Big|_{v=0}\mathrm{d}t.
\end{aligned}$$

Plugging this equality into the law of motion of $w_{\mathcal{P}t}$ (30), and using the definition of cumulants (31), we obtain

$$\begin{aligned}
\frac{\mathrm{d}w_{\mathcal{P}t}}{w_{\mathcal{P}t}}&=\mathrm{E}^{g_t(\cdot)\cdot}\left[\mu_t(w)|w\geq q_t\right]\mathrm{d}t+\frac{1}{pw_{\mathcal{P}t}}\sum_{n=2}^{\infty}\frac{1}{n!}\partial_v^n\left(\int_{q_te^{-v}}^{q_t}(e^vw-q_t)\kappa_{nt}(w)g_t(w)\,\mathrm{d}w\right)\Big|_{v=0}\mathrm{d}t\\
&=\mathrm{E}^{g_t(\cdot)\cdot}\left[\mu_t(w)|w\geq q_t\right]\mathrm{d}t+\frac{1}{pw_{\mathcal{P}t}}\sum_{n=2}^{\infty}\frac{1}{n!}\partial_v^{n-1}\left(\int_{q_te^{-v}}^{q_t}e^vw\kappa_{nt}(w)g_t(w)\,\mathrm{d}w\right)\Big|_{v=0}\mathrm{d}t\\
&=\mathrm{E}^{g_t(\cdot)\cdot}\left[\mu_t(w)|w\geq q_t\right]\mathrm{d}t+\frac{q_t}{pw_{\mathcal{P}t}}\sum_{n=2}^{\infty}\frac{1}{n!}\sum_{m=0}^{n-2}(-w\partial_w)^m\left(\kappa_{nt}(w)wg_t(w)\right)|_{w=q_t}\,\mathrm{d}t.
\end{aligned}$$

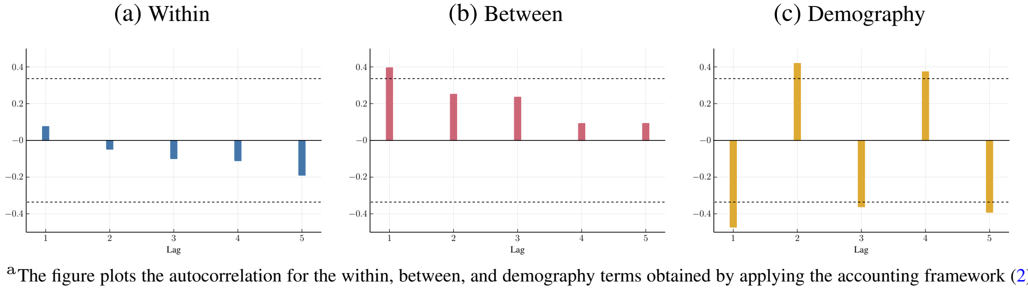
Q.E.D.

APPENDIX D: EMPIRICAL APPENDIX

D.1. *Autocorrelogram*

Figure D.1 plots the serial correlation of the within, between, and demography terms. Each term has different patterns. The within term is serially uncorrelated, which reflects the fact that stock market returns are approximately uncorrelated over time. The between term is positively correlated over time, with an AR(1) structure, which reflects the fact that the dispersion of wealth growth among top households tends to be serially correlated over time (as seen in Figure 4a). Finally, the demography term is negatively correlated at the one year horizon, with a MA(1) structure, which reflects the fact that the fraction of top households that die in a given year tends to be negativelly correlated over time. Intuitively, a higher number of death in a given year implies fewer death in the following year, as the remaining population is younger.

FIGURE D.1.—Autocorrelogram for the Within, Between and Demography Terms in the Forbes 400 [a]



[a]The figure plots the autocorrelation for the within, between, and demography terms obtained by applying the accounting framework (2) on the growth of the Forbes 400 wealth share. The dashed lines represent the 95% confidence bands corresponding to the null hypothesis of no correlation. Data are from *Forbes* and the Financial Accounts of the United States.

D.2. *Measuring the Wealth of Drop-Offs*

The decomposition in Section 2 requires knowing the wealth of households that drop off the top percentile. However, before 2012, *Forbes* only rarely reported the wealth of individuals who dropped out from the top 400. First, 70% of households that drop off the top percentile actually stay in the Forbes 400. Indeed, the top percentile used in this paper is composed of only 264 households in 1983 (it was chosen so that, with population growth, it includes 400 households in 2017). Because wealth is so concentrated in the top, there is usually a great difference between the last individual in this top percentile and the wealth of the last individual in the top 400. Therefore, most households that drop off this top percentile stay in the top 400.

I now focus on the remaining 30% of households that drop off the Forbes 400. Formally, the problem boils down to estimating the average of a variable (the wealth growth of top households) that is left censored. In this particular setting, the Kaplan and Meier (1958) estimator gives tight bounds to estimate this quantity. The idea is to estimate the average growth rate of drop-offs using the observed negative growth rates of households in the top percentile. The identifying assumption is that the distribution of growth rates is homogeneous for households within the top percentile.

More precisely, the Kaplan and Meier (1958) method is to first estimate the survival function, that is, in my setting the probability that wealth growth is lower than a certain threshold $\mathrm{P}((w_{it+1} - w_{it})/w_{it} \leq x)$. This survival function can then be used to estimate the expectation of wealth growth conditional on being lower than a certain threshold, that is, $\mathrm{E}[(w_{it+1} - w_{it})/w_{it} | (w_{it+1} - w_{it})/w_{it} \leq x]$.

I check the validity of this imputation method by focusing on years where *Forbes* reports the wealth of drop-offs (i.e., 2012-2017). In these years, I compare the result obtained from the estimated method and the result obtained using the real wealth of drop-offs. Table D.I shows the average return of these drop-offs using the imputed method and the actual data reported by wealth. The estimates differ by only 2 percentage points, on average (-26% versus -28%). The fact that the Kaplan-Meier estimator gives such a good result is intuitive: because wealth is so concentrated, households at the very top of the wealth distribution hold ten times more wealth than the households at the percentile threshold, and therefore I actually observe a large part of the distribution of negative wealth shocks.

The last four columns report the estimates for the within and between terms using imputed and real data. The estimates differ by 5 basis points. The bias is small because, as discussed above, the Kaplan-Meier method provides accurate estimates of the wealth growth of imputed households. Moreover, the wealth share represented by the imputed households is small to begin with.

I use the same method to impute the second, third, and fourth order of wealth to obtain the estimate of standard deviation, skewness, and excess kurtosis reported in Table IIa.

TABLE D.I

EFFECT OF USING IMPUTED VERSUS REPORTED WEALTH OF DROP-OFFS ON DECOMPOSITION[a]

Year	Average Wealth Growth of Drop-Offs (%)		Within (%)		Between (%)	
	Imputed	Actual	Imputed	Actual	Imputed	Actual
2011-2012	−33.56	−36.48	7.75	7.70	1.22	1.26
2012-2013	−27.34	−29.82	6.77	6.73	1.13	1.17
2013-2014	−23.82	−19.35	2.36	2.44	1.69	1.61
2014-2015	−26.44	−30.28	−2.99	−3.08	1.34	1.42
2015-2016	−26.46	−32.53	3.29	3.15	1.33	1.46
2016-2017	−16.97	−19.14	1.49	1.44	1.61	1.65
2011-2017	−25.93	−28.22	3.05	3.00	1.38	1.43

[a]The table compares the within and the between terms (defined in (2)) using imputed data and actual data about the wealth of drop-offs. Starting from 2011, *Forbes* systematically reports the wealth of all drop-offs. For these years, one can compare the results of the accounting framework using this reported data as opposed to the imputation using Kaplan and Meier (1958). All terms in percentage. Data are from *Forbes* and the Financial Accounts of the United States.

D.3. *Assessing the Importance of Measurement Error in Wealth*

How does transitory measurement errors in wealth affect the accounting framework? Suppose that wealth follows a simple diffusion process with drift μ and volatility ν. Moreover, suppose that wealth is only observed every year, and only with some errors; i.e., we only observe $\hat{w}_{it} = w_{it} e^{\epsilon_{it}}$, where ϵ_{it} is an i.i.d process with $\mathrm{E}[e^{\epsilon_{it}}] = 1$.

The log change in observed wealth between t and $t+1$ can be written as

$$\log\left(\frac{\hat{w}_{it+1}}{\hat{w}_{it}}\right) = \mu + \nu(B_{it+1} - B_{it}) + \epsilon_{it+1} - \epsilon_{it}.$$

Denote ρ the slope coefficient in a regression on observed wealth growth on its lagged value. We can relate ρ to the variance of measurement errors, $\mathrm{Var}(\epsilon_{it})$:

$$\rho = \frac{\mathrm{cov}(\log(\hat{w}_{it+1}/\hat{w}_{it}), \log(\hat{w}_{it}/\hat{w}_{it-1}))}{\mathrm{Var}(\log(\hat{w}_{it}/\hat{w}_{it-1}))}$$

$$= \frac{\text{cov}(\mu + \nu(B_{it+1} - B_{it}) + \epsilon_{it+1} - \epsilon_{it}, \mu + \nu(B_{it} - B_{it-1}) + \epsilon_{it} - \epsilon_{it-1})}{\text{Var}(\mu + \nu(B_{it} - B_{it-1}) + \epsilon_{it} - \epsilon_{it-1})}$$

$$= -\frac{\text{Var}(\epsilon_{it})}{\nu^2 + 2\text{Var}(\epsilon_{it})}$$

This equation makes it possible to recover the variance of measurement errors from the serial correlation of log wealth growth. More precisely, Table D.II reports an estimate for $\rho \approx -0.01$, which implies $\text{Var}(\epsilon_{it})/\nu^2 \approx 1\%$.

Denote $\widehat{\text{Between}}$ the between term obtained by applying the accounting framework on $\hat{w}_{it}$ between t and $t+1$. Over a short time period, Proposition 2 says that the relative bias in the between term is equal to the relative bias in the variance of log wealth growth:

$$\frac{\widehat{\text{Between}} - \text{Between}}{\text{Between}} \approx 2\text{Var}(\epsilon_{it})/\nu^2$$

Overall, this exercise suggests that the relative bias in the between term is around 2%, which is very small.

TABLE D.II

SERIAL CORRELATION OF WEALTH GROWTH IN THE FORBES 400[a]

	Wealth Growth
	(1)
Lagged Wealth Growth	–0.01
	(0.01)
Constant	0.05
	(0.00)
R^2	0.18
Period	1983-2016
FE	Individual
N	11,392

[a]The table shows the result of regressing future wealth growth on current wealth growth with individual fixed effects; that is, denoting w_{it} the wealth of household i at time t:

$$\log\left(\frac{w_{it+1}}{w_{it}}\right) = \alpha_i + \beta \log\left(\frac{w_{it}}{w_{it-1}}\right) + \epsilon_{it+1}.$$

Estimation is done via OLS. Standard errors are in parentheses and are estimated using Newey-West with three lags. Data are from *Forbes*.

APPENDIX E: DECOMPOSING THE GROWTH OF BILLIONAIRE'S SHARES IN RUSSIA AND CHINA

I now compare the accounting decomposition obtained for the U.S. with other countries, using *Forbes*'s list of international billionaires starting in 1987. One difficulty is that the number of billionaires in each country tends to be much smaller than in the U.S. Therefore, I restrict myself to Russia and China, the two countries with the highest count of billionaires in 2010 outside the U.S.[56] I consider the wealth share of the percentile composed of 50 billionaires in

[56]For the sake of the exercise, I group together China, Hong Kong, and Taiwan.

2010. Since China counts less than 50 billionaires before 2010, I only look at the growth of this wealth share in the 2010–2017 period.[57]

Another concern is that data quality may be lower for these two countries compared to the U.S. In particular, the addition of a billionaire to the list may reflect the fact that Forbes becomes more familiar with the country over time, rather than a real change in the underlying wealth distribution. To address this concern, I manually check that each new Chinese or Russian billionaire can be traced back to a particular event, such as a successful IPO or a high stock market return; otherwise, I remove the household from the sample. Finally, data on per capita household wealth in Russia and China are taken from the World Wealth and Income Database.

Table E.I shows the result of the accounting decomposition for these two countries as well as the results of the U.S. for the same time period. Figure E.1a plots the average within, between, and demography terms in each country over time as well as their standard errors. Wealth inequality increased in both countries: the yearly growth rate of the top share is 1.0% in China and 8.0% in Russia. The average within term has a large standard error, which reflects the short time window for the decomposition as well as the small number of households in each group. In contrast, the between term is much more precisely estimated. As a consequence, in the rest of the analysis, I focus on comparing the between term across countries.

The between term averages to 1.1% in Russia over the last ten years, which is roughly similar to the average between term in the U.S. over the same time period. In contrast, the between term averages to 4.2% in China, which is three times higher than the average between term in the U.S. To better understand this difference, I compare the model-implied between term $1/2 \cdot g_t(q_t)q_t^2/(pw_{\mathcal{P}t}) \cdot \nu_t(q_t)^2$ in China, using the same methodology as the U.S. (see Figure E.1b). This allows me to decompose the change between the between term in China and in the U.S. as a sum of two factors:

$$\underbrace{\Delta\left\langle \frac{1}{2}\frac{g_t(q_t)q_t^2}{pw_{\mathcal{P}t}}\nu_t(q_t)^2 \right\rangle}_{2.7\%} = \underbrace{\left\langle \frac{1}{2}\frac{g_t(q_t)q_t^2}{pw_{\mathcal{P}t}} \right\rangle \Delta\left(\nu_t(q_t)^2\right)}_{1.3\%} + \underbrace{\langle \nu_t(q_t)^2\rangle \Delta\left(\frac{1}{2}\frac{g_t(q_t)q_t^2}{pw_{\mathcal{P}t}}\right)}_{1.4\%},$$

where Δx (resp. $\langle x\rangle$) denotes the difference (resp. average) of variable x between China and the U.S. I find that the relatively high level of the between term in China comes from the combination of two distinct forces. First, the standard deviation of (log) wealth growth among top households is much higher in China compared to the U.S. Second, wealth inequality in China is much lower to begin with, which makes it easier for households with high realized growth rates to displace existing fortunes at the top.

APPENDIX F: DECOMPOSING THE GROWTH OF TOP INCOME SHARES

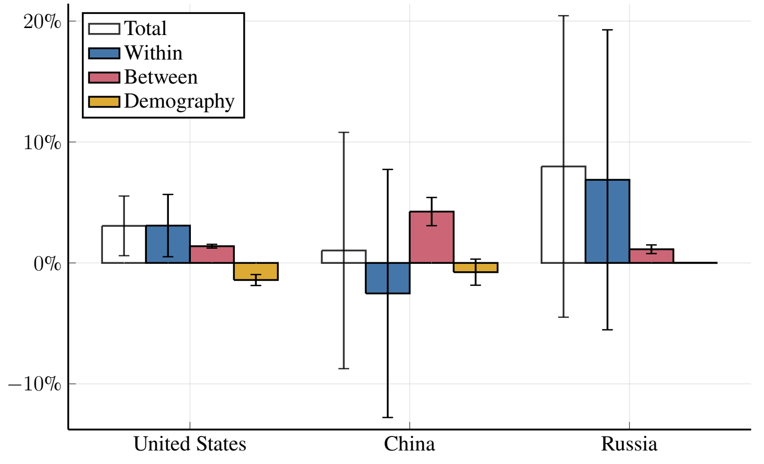
I now apply the accounting framework to decompose the growth of top income shares in the U.S. The key difference with wealth is that income is a flow, not a stock. In particular, transitory shocks play a much larger role in driving the dynamics of income in the short run, even though they only play a small role in shaping the income distribution.[58] The situation is similar to the measurement error problem discussed in Appendix D: while we would like to do the decomposition on the permanent component of labor income, we only observe a noisy version of it.

[57]The evolution of wealth inequality in China and Russia has been also discussed in Novokmet et al. (2018) and Piketty et al. (2019).

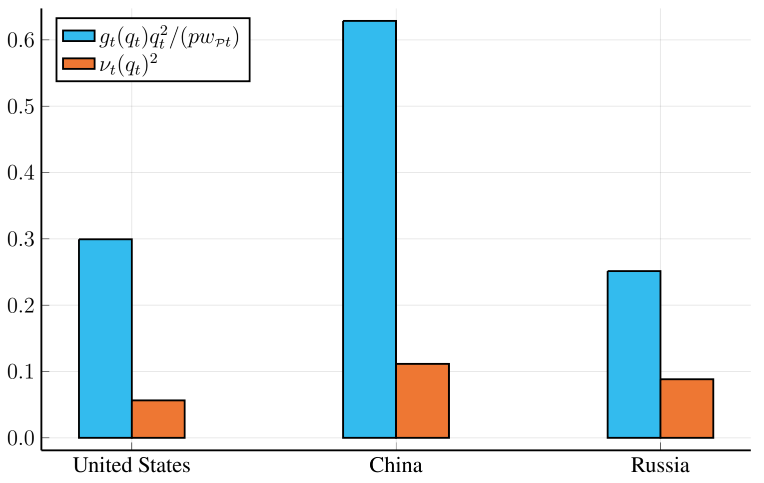
[58]See Heathcote et al. (2010) and Blundell et al. (2008).

FIGURE E.1.—Decomposing the Annual Growth Rate of Billionaires' Shares (2010-2017)[a]

(a) Within, Between, and Demography Terms



(b) Estimates of $g_t(q_t)q_t^2/(pw_{\mathcal{P}t})$ and $\nu_t(q_t)^2$



[a] Figure E.1a plots the average annual growth rate of the billionaires' shares in the U.S., Russia, and China, as well as the within, between, and demography terms defined in the accounting framework (2). The figure also plots the 95% confidence-interval for the average of each quantity over the time period. Figure E.1b plots the average estimates of $g_t(q_t)q_t^2/(pw_{\mathcal{P}t})$ and $\nu_t(q_t)^2$ for these three countries. Data are from *Forbes* and the World Wealth and Income Database.

One simple way to adapt the accounting framework to top income shares is to apply the decomposition on a more "permanent" measure of income: the average income received by each individual averaged over a three year period.

I use the IRS public use panel files created by the Statistics of Income Division from 1979 to 1990. This dataset includes a random sub-sample of taxpayers who can be followed over time between 1979 and 1990. Following Piketty and Saez (2003), I construct a comprehensive

TABLE E.I

DECOMPOSING THE GROWTH OF THE BILLIONAIRES' WEALTH SHARE[a]

(a) Summary

	Growth of Top Percentile Wealth Share			
	Total	Within	Between	Demography
United States	3.1	3.1	1.4	−1.4
China	1.0	−2.5	4.2	−0.8
Russia	8.0	6.9	1.1	0.0

(b) Within

	Within		
	Total	Average Wealth Growth in Top Percentile	−Growth Country Wealth per Capita
United States	3.1	8.1	−4.8
China	−2.5	9.1	−11.9
Russia	6.9	2.9	3.8

(c) Between

	Between						
	Total	Inflow			Outflow		
		Total	$n_{\mathcal{I}}$	$\frac{w_{\mathcal{I},1}-q_1}{w_{\mathcal{P}_0,0}}$	**Total**	$n_{\mathcal{O}}$	$\frac{q_1-w_{\mathcal{O},1}}{w_{\mathcal{P}_0,0}}$
United States	1.4	1.1	8.2	14	0.2	5.9	4
China	4.2	2.8	17.6	17	1.4	16.7	9
Russia	1.1	0.7	10.6	6	0.5	10.6	4

(d) Demography

	Demography									
	Total	Birth			Death			Population Growth		
		Total	$n_{\mathcal{B}}$	$\frac{w_{\mathcal{B},1}-q_1}{w_{\mathcal{P}_0,0}}$	**Total**	$n_{\mathcal{D}}$	$\frac{q_1-\frac{w_{\mathcal{P}_0\setminus\mathcal{D},1}}{w_{\mathcal{P}_0\setminus\mathcal{D},0}}w_{\mathcal{D},0}}{w_{\mathcal{P}_0,0}}$	**Total**	$1-n_{\mathcal{P}_0}$	$\frac{q_1-\frac{w_{\mathcal{P}_0\setminus\mathcal{D},1}}{w_{\mathcal{P}_0\setminus\mathcal{D},0}}w_{\mathcal{P}_0,0}}{w_{\mathcal{P}_0,0}}$
United States	−1.4	0.3	0.3	78	−0.9	1.6	−56	−0.8	1.0	−74
China	−0.8	0.1	0.6	21	−0.7	1.1	−63	−0.2	0.3	−62
Russia	0.0	0.0	0.0	—	0.0	0.0	—	0.0	0.0	—

[a] Table E.I reports the geometric mean of the growth rate of Billionaires' shares as well its decomposition into a within, between, and demography terms as defined in accounting framework (2). The within term is further decomposed into the growth of the average wealth of households initially in the top percentile and the growth of the per capital wealth in each country (both expressed in real dollars). All terms in percentage. Data are from *Forbes*, the Bureau of Economic Analysis, and the World Wealth and Income Database.

measure of earnings, which includes wages, salary, and entrepreneurial income.[59] I focus on tax filers who file jointly — note that, in this context, birth and death represent taxpayers who start or stop filing jointly. The sample contains 5,523 distinct taxpayers. Finally, I replace values of income below the bottom 10% (in a given year) by the bottom 10% quantile. While this

[59]More precisely, I construct earnings as adjusted growth income minus capital gain, dividend income, interest income, rental income, and royalties.

winsorizing does not play an important role for the results of the accounting decomposition, it helps making the distribution of log income growth closer to a normal distribution, and, therefore, to obtain a better approximation by the diffusion model obtained in Section 3.

As mentioned earlier, I want to focus on the permanent component of income. To do so, I average income over three-year periods before applying the accounting on the sample. I then annualize each term to obtain yearly terms (i.e., dividing them by three). I find similar results using a larger time period such as six years, which suggests that three years is enough to focus on the permanent component of income.

Table F.I shows the result of the accounting decomposition for the top 100%, top 10%, and top 2% (due to fact I only observe a sample of taxpayers, I cannot delve into smaller top percentiles — the top 2% only represents 100 taxpayers). Note that, by definition, the growth of top income share and the between term for the top 100% are equal to zero.

I find that the annual growth rate of top income shares during the time period for the top 2% is 4.0%. This number can be decomposed as follows: a within term of 0.3%, a between term of 4.2%, and a demography term of -0.5%. In other words, the between term plays an even larger role compared to top wealth shares: the between term accounts for the whole increase in top income inequality.

Figure F.1 plots the result of the decomposition over all percentiles. While the within term and the demography term are relatively stable across the income distribution, the between term gradually increases with top percentiles. As a result, the increase of the between term with respect to percentiles lines up with the increase of the top income share, suggesting again that the between term has been a key driver of the rise in top income shares during the time period.

To understand why the between term increases with top percentiles, Figure F.1a plots the contributions of $g_t(q_t)q_t^2/(pw_{\mathcal{P}t})$ and $\nu_t(q_t)^2$ as a function of the top percentile. The standard deviation of log income follows a U-shape, as in Guvenen et al. (2015). However, $g_t(q_t)q_t^2/(pw_{\mathcal{P}t})$ increases almost monotonically with the top percentile, except at the very top. Overall, I find that the model-implied between term increases monotonically with top percentiles.

TABLE F.I

DECOMPOSING THE GROWTH OF TOP INCOME SHARES (1979-1990)[a]

(a) Summary

	Growth of Top Percentile Wealth Share			
	Total	Within	Between	Demography
Top 100%	0.0	–0.1	0.0	0.1
Top 10%	2.2	–0.0	2.3	–0.1
Top 2%	4.0	0.3	4.2	–0.5

(b) Within

	Within		
	Total	Average Income Growth in Top Percentile	−Growth of U.S. Income per Capita
Top 100%	–0.1	2.5	–2.6
Top 10%	–0.0	2.6	–2.6
Top 2%	0.3	2.9	–2.6

(c) Between

	Between						
	Total	Inflow			Outflow		
		Total	$n_{\mathcal{I}}$	$\frac{w_{\mathcal{I},1}-q_1}{w_{\mathcal{P}_0,0}}$	**Total**	$n_{\mathcal{O}}$	$\frac{q_1-w_{\mathcal{O},1}}{w_{\mathcal{P}_0,0}}$
Top 100%	0.0	0.0	0.0	—	0.0	0.0	—
Top 10%	2.3	1.2	7.3	16	1.2	7.4	16
Top 2%	4.2	2.5	9.2	26	1.8	8.8	21

(d) Demography

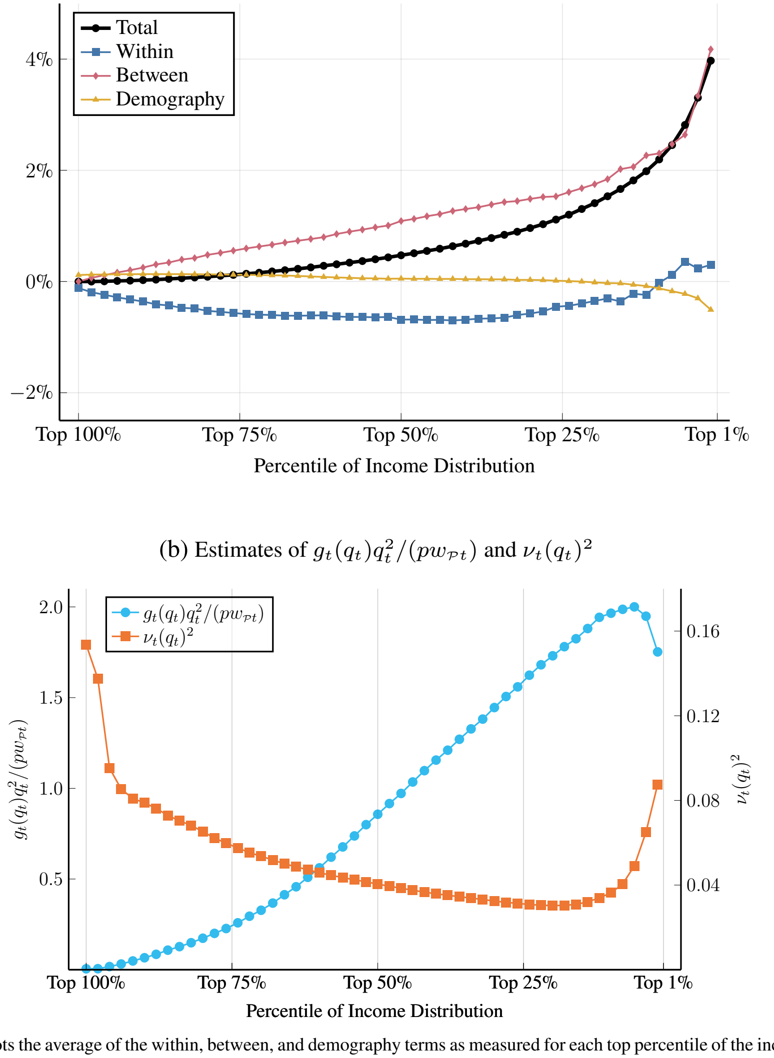
	Demography									
	Total	Birth			Death			Population Growth		
		Total	$n_{\mathcal{B}}$	$\frac{w_{\mathcal{B},1}-q_1}{w_{\mathcal{P}_0,0}}$	**Total**	$n_{\mathcal{D}}$	$\frac{q_1-\frac{w_{\mathcal{P}_0\setminus\mathcal{D},1}}{w_{\mathcal{P}_0\setminus\mathcal{D},0}}w_{\mathcal{D},0}}{w_{\mathcal{P}_0,0}}$	**Total**	$1-n_{\mathcal{P}_0}$	$\frac{q_1-\frac{w_{\mathcal{P}_0\setminus\mathcal{D},1}}{w_{\mathcal{P}_0\setminus\mathcal{D},0}}w_{\mathcal{P}_0,0}}{w_{\mathcal{P}_0,0}}$
Top 100%	0.1	3.8	5.1	75	–3.0	4.2	–72	–0.9	1.0	–91
Top 10%	–0.1	1.1	3.5	32	–1.0	2.4	–38	–0.3	1.0	–35
Top 2%	–0.5	0.9	3.0	33	–1.2	2.6	–37	–0.3	0.9	–30

[a] Table F.I reports the average of the annual growth rate for the top 100%, 10%, and 2% income shares, as well as the within, between, and demography terms defined in the accounting framework (2). All terms in percentage. Data are from the IRS public use panel files.

FIGURE F.1.—Decomposing the Growth of Top Income Shares (1979-1990)[a]

(a) Within, Between, and Demography Terms

(b) Estimates of $g_t(q_t)q_t^2/(pw_{\mathcal{P}t})$ and $\nu_t(q_t)^2$



[a] Figure F.1a plots the average of the within, between, and demography terms as measured for each top percentile of the income distribution. Figure F.1b plots the average estimates of $g_t(q_t)q_t^2/(pw_{\mathcal{P}t})$ and $\nu_t(q_t)^2$ for each top percentile of the income distribution. Data are from the IRS public use panel files.