Online Appendix for "Rise and Decline of General Laws of Capitalism"

In this Appendix, we discuss the core theoretical claims of Piketty's *Capital in the 21st Century*, in an effort to clarify the relationship between $r - g$ and inequality. The emphasis will be on four issues: (1) what types of models and economic forces lead to a divergence of inequality when $r > g$; (2) the role of social mobility in this process; (3) what types of models lead instead to a relationship between $r - g$ and the (right) tail of the stationary distribution of income and wealth; (4) how does $r - g$ respond to policies and capital accumulation.

Divergence of Inequality when $r - g > 0$ (without Social Mobility)

The first possible reading of the theoretical core of *Capital in the 21st Century* is that if $r - g$ is positive (or sufficiently large) it will lead to a divergence of wealth between the very rich and the rest of population. The approach of the book here builds on ideas proposed by Nicholas Kaldor, in particular, Kaldor (1955). As we will see, this model gives a formalization of the various intuitions and statements made in *Capital in the 21st Century* in a rather straightforward manner, but also shows what the limitations of some of these intuitions and claims are.

The prototypical Kaldor-type economy consists of "capitalists" and workers (and no land), and an important dimension of inequality is between these two groups and is fueled by the assumption that capitalists have a high saving rate (and workers have a saving rate of zero), and all of the income of the capitalists come from capital. As we will see, there is no need to assume that workers do not have any capital income; it is sufficient to allow different saving rates between these two groups.

Suppose that the economy comprises a single good, so that there is no relative price for installed capital (relative to final output and consumption). We also focus on a continuous-time economy for notational convenience. Let us denote the capital stock held by capitalists by K_C. For future reference, we also denote the fraction of capitalists in the population by m, and thus the fraction of workers is $1 - m$, and without loss of generality, we take these to be the numbers of capitalists and workers (thus normalizing total population to 1). For now, there is no social mobility between capitalists and workers, but we will relax this below.

Since all of the income of the capitalists comes from capital, their total income is simply given by the capital stock times the rental price of capital. Assuming that capital depreciates at the rate δ and the interest rate is r, total income accruing to capitalists can be written as

$$I_C = (r + \delta)K_C, \tag{A1}$$

where we are suppressing time indices.[9]

[9] Piketty specifies everything, including the saving rate in net of depreciation units. But as Krusell and

Now supposing that capitalists have a constant saving rate of s_C out of their income, the evolution of the capital stock held by capitalists can be written as

$$\begin{aligned} \dot{K}_C &= s_C I_C - \delta K_C \\ &= [s_C(r+\delta) - \delta]K_C, \end{aligned}$$

where the first line simply uses the fact that a constant fraction s_C of capitalists' income, I_C, is saved, but then a fraction δ of their existing capital stock depreciates. The second line simply substitutes for I_C from (A1).

To obtain the growth rate of capitalists' income, we also need to know how the interest rate varies over time. In particular, the growth rate of capitalists' income can be obtained by differentiating (A1) with respect to time as

$$\begin{aligned} g_C^I &= \frac{\dot{K}_C}{K_C} + \frac{\dot{r}}{r+\delta} \\ &= s_C(r+\delta) - \delta + \frac{\dot{r}}{r+\delta}. \end{aligned}$$

Now returning to workers, their income is

$$\begin{aligned} I_W &= (r+\delta)K_W + wL \\ &= (r+\delta)K_W + Y - (r+\delta)(K_C + K_W) \\ &= Y - (r+\delta)K_C, \end{aligned}$$

where K_W is the capital stock held by workers, w the real wage, L total employment and where the second line simply uses the fact that labor income is equal to national income minus capital income. Then, the growth rate of the income of workers can be obtained by straightforward differentiation with respect to time and by rearranging terms using the expression for the income of the capitalists from (A1):

$$g_W^I = \frac{\frac{\dot{Y}}{Y} - \frac{\dot{K}_C}{K_C}\frac{I_C}{Y} - \frac{\dot{r}}{r+\delta}\frac{I_C}{Y}}{1 - \frac{I_C}{Y}}.$$

One advantage of this expression is that it is written without reference to the saving rate of workers, s_W, because of the national income accounting identity. But this is also a disadvantage, because, as we discuss below, it masks that comparisons of r to g are implicitly changing the growth of labor income and the saving rate of workers.

Denote the fraction of national income accruing to capitalists by ϕ $(= I_C/Y)$. If capitalists correspond to the richest 1 percent in the population, then ϕ is the top 1 percent share measure

Smith (2014) emphasize, this is a difficult assumption to motivate and leads to the unpleasant and untenable implication of all of national income being saved at low growth rates. In light of this, it is more appropriate to think of Piketty's results as being supported by assuming that $\delta \approx 0$.

used extensively by Piketty. Using this definition and denoting the growth rate of GDP (and national income) by g, we can then write

$$g_W^I = \frac{g - [s_C(r+\delta) - \delta]\phi - \frac{\dot{r}}{r+\delta}\phi}{1-\phi}.$$

Let us now compare this to the growth rate of the income of the capitalists. A simple rearrangement gives that

$$\begin{aligned} g_C^I &> g_W^I \text{ if and only if} \\ s_C(r+\delta) &> g + \delta - \frac{\dot{r}}{r+\delta}. \end{aligned} \tag{A2}$$

This expression thus states that there will be divergence between the incomes of the capitalists and the workers when (A2) holds.[10] Note, however, that this sort of divergence, by definition, must be temporary, because if capitalists' incomes are growing faster than the rest of the population, at some point they will make up the entire income of the economy.[11]

It is now straightforward to observe that the claim in *Capital in the 21st Century* about $r - g > 0$ leading to expanding inequality will hold under two additional conditions:

1. $s_C \simeq 1$, which would follow if the very rich save a very large fraction of their incomes. In practice, though the very rich save more of their incomes than the poor, s_C is significantly less than 1, especially once one takes into account charitable contributions and donations by the very rich.

2. $\dot{r} = 0$, so that the interest rate is constant. Here, as discussed in the text, much of growth theory suggests that the interest rate is quite responsive to changes in the capital stock (and other factors of production as well as technology).

Under these two assumptions, (A2) boils down to

$$g_C^I > g_W^I \text{ if and only if } r > g,$$

as asserted by Piketty. However, (A2) also makes it clear that without the two simplifying assumptions above, the evolution of top inequality depends on the saving rate and changes in the interest rate as well as $r > g$.

[10] See also Homburg (2014) for an explanation for why $r - g$ does not translate to divergence in overlapping generations models.

[11] In particular, when (A2) holds for an extended period of time, then all of national income will be in terms of capital income, so it is impossible for $r > g$ and thus for (A2) to be maintained forever.

Divergence of Inequality with Social Mobility

The simple Kaldor-type model presented in the previous subsection enables us to present a transparent illustration of how social mobility affects inequality. We will now show that even under the assumptions enumerated above, modest amounts of social mobility can significantly change the conclusions. Though the United States is not one of the highest social mobility countries in the world, it still has a fairly sizable likelihood that those at the top of income distribution will lose their position, and as mentioned in the text, recent evidence by Chetty, Hendren, Kline and Saez (2014a,b) suggests that this rate of social mobility has not declined over time, even though overall inequality has increased sharply.

Let us now incorporate the possibility of social mobility into this simple framework. To simplify the exposition, let us suppose that $\delta = \dot{r} = 0$ for this part of the analysis.

We model social mobility as follows. We assume that at some flow rate ν, a capitalist is hit by a random shock and becomes a worker, inheriting the worker's labor income process and saving rate. At this point, he (or she) of course maintains his current income, but from then on his income dynamics follows those of other workers. Simultaneously, a worker becomes a capitalist (also at the flow rate ν), keeping the fraction of capitalists in the population constant at $m \in (0, 1)$.

We can now write the dynamics of the total income of capitalists as

$$\dot{I}_C = s_C r I_C - \nu \left[\frac{I_C}{m} - \frac{I_W}{1-m} \right], \tag{A3}$$

where we are exploiting the fact that, on average, a capitalist leaving the capitalist class has income I_C/m (total capitalists' income divided by the measure of capitalists), and a worker entering the capitalist class has, on average, income $I_W/(1-m)$. This significantly facilitates the characterization of inequality between capitalists and workers (though the determination of the exact distribution of income is more complicated because of the slow growth dynamics of the income of individuals that change economic class).[12]

Now rearranging (A3), we obtain

$$\begin{aligned} g_C^I &= s_C r - \nu \left[\frac{1}{m} - \frac{1}{1-m} \frac{I_W}{I_C} \right] \\ &= s_C r - \nu \left[\frac{1}{m} - \frac{1}{1-m} \frac{1-\phi}{\phi} \right]. \end{aligned}$$

With a similar reasoning, the growth rate of the total income of workers is

$$g_W^I = \frac{g - s_C r \phi}{1-\phi} + \nu_W \left[\frac{1}{m} \frac{\phi}{1-\phi} - \frac{1}{1-m} \right].$$

[12] This also means that the comparison of the incomes of capitalists and workers in this world with social mobility is only an approximation to the top 1 percent inequality measures (even when capitalists make up 1 percent of the population), because workers who become capitalists will join the top 1 percent only slowly.

Combining these expressions and rearranging terms, we can write

$$
\begin{aligned}
g_C^I &> g_W^I \text{ if and only if} \\
s_C r - g &> \nu \frac{\phi - m}{\phi m(1-m)},
\end{aligned} \tag{A4}
$$

where the term on the right-hand side is strictly positive in view of the fact that $\phi > m$ (i.e., the share of top 1 percent in national income is greater than 1%). This expression thus shows that even when $s_C r - g > 0$ (or, fortiori, when $r - g > 0$), it does not follow that inequality between capitalists and workers will increase. Whether it does will depend on the extent of social mobility. In fact, the quantitative implications of social mobility could be quite substantial as we next illustrate.

From Chetty, Hendren, Kline and Saez's data, the likelihood that a child with parents in the top 1 percent will be in the top 1 percent is 9.6%.[13] If we take the gap between generations to be about 30 years, this implies an annual rate of exiting the top 1 percent approximately equal to 0.075 (7.5%). There are many reasons why this may be an overestimate, including the fact that children are typically younger when their incomes are measured and also that in practice, families exiting the top 1 percent tend to remain at the very top of the income distribution (rather than follow the income dynamics of a typical worker as in the simple model here). But there are also reasons for underestimation, including the fact that within-generation transitions in and out of the top 1 percent are being ignored. For our illustrative exercise, we take this number as a benchmark (without any attempt to correct it for these possible concerns). This number corresponds to ν/m in our model (the probability that a given capitalist is hit by a shock and becomes a worker), so we take $\nu = 0.00075$. Using the top 1 percent's share as 20%, we can compute that the right-hand side of (A4) is approximately 0.072 (72%). This implies that for the left-hand side to exceed the right-hand side, the interest rate would have to be very high. For example, with a saving rate of 50% and a growth rate of 1%, we would need the interest rate to be greater than 15%. Alternatively, if we use the top 10 percent so as to reduce exits that may be caused by measurement error, the equivalent number from Chetty, Hendren, Kline (2014) is 26%, implying an annual exit rate equal to 4.4%. Using a share of 45% of income for the top 10 percent, the right-and side of (A4) can be computed as 0.038, again making it very difficult for realistic values of $r - g$ to create a natural and powerful force for the top inequality to increase. For example, using again a saving rate of 50% and a growth rate of 1%, the interest rate would need to be over 8.5%. We therefore conclude that incorporating social mobility greatly reduces any "fundamental force" that may have existed from $r - g$ towards mechanically greater inequality at the top of the distribution.

[13] We thank Nathan Hendren for providing us with the data.

$r - g$ and the Stationary Distribution of Income and Wealth

As discussed in the text, *Capital in the 21st Century* sometimes posits a relationship between $r - g$ and the stationary distribution of wealth instead of the relationship between $r - g$ and divergence of incomes and wealth. Empirically the Pareto distribution (with distribution function $1 - \Gamma a^{-\lambda}$ with Pareto shape coefficient $\lambda \geq 1$) appears to be a good approximation to the tail of the income and wealth distributions. For this reason, existing models have focused on stochastic processes for wealth accumulation that generate a Pareto distribution or distributions for which the right tail can be approximated by the Pareto form. Such models have a long history in economics, and are discussed in the context of the issues raised in *Capital in the 21st Century* in Jones (2014), and we refer the reader to this paper for more extensive references. Some recent papers that derive Pareto wealth distributions include Benhabib, Bisin and Zhu (2011), Aoki and Nirei (2013) and Jones and Kim (2014).

In this part of the appendix, we show that a Pareto tail in the wealth distribution emerges from certain classes of models, and will, under some conditions, correspond to greater top inequality when $r - g$ is higher, but we also highlight why these models are often not a good approximation to the type of top inequality we observe in the data and/or rely on implausible assumptions.

To give the basic idea, consider an economy consisting of a continuum of measure 1 of heterogeneous individuals. Suppose that each individual i is infinitely lived and consumes a constant fraction β of her wealth, A_{it}. She has a stochastic (possibly serially correlated) labor income Z_{it} (with $\mathbb{E}Z_{it} \in (0, \infty)$ and finite variance), and has a stochastic rate of return equal to $r + \varepsilon_{it}$, where ε_{it} is a stochastic, return term that is also possibly serially correlated (with the unconditional mean $\mathbb{E}\varepsilon_{it}$ equal to zero as a normalization). Thus, the law of motion of the assets of individual i is given by

$$A_{it+1} = (1 + r - \beta + \varepsilon_{it})A_{it} + Z_{it}.$$

Dividing both sides of this equation by GDP (also average income per capita), Y_t, we obtain

$$\tilde{a}_{it+1} = \frac{1 + r - \beta + \varepsilon_{it}}{1 + g}\tilde{a}_{it} + \tilde{z}_{it},$$

where $\tilde{a}_{it} \equiv A_{it}/Y_t$ and $\tilde{z}_{it} \equiv Z_{it}/Y_t$. A further normalization is also useful. Suppose that $\tilde{a}_{it}$ converges to a stationary distribution (we verify this below). Then let $\mathbb{E}\tilde{a}$ be the average (expected) value of $\tilde{a}_{it}$ in the stationary distribution. Then dividing both sides of this equation by $\mathbb{E}\tilde{a}$, we obtain

$$a_{it+1} = \frac{1 + r - \beta + \varepsilon_{it}}{1 + g}a_{it} + z_{it}, \tag{A5}$$

where $a_{it} \equiv \tilde{a}_{it}/\mathbb{E}\tilde{a}$ and $z_{it} \equiv \tilde{z}_{it}/\mathbb{E}\tilde{a}$, and of course $\mathbb{E}a_{it+1} = \mathbb{E}a_{it} = 1$ in the stationary distribution. This also implies that $\mathbb{E}z_{it} \in (0, 1)$.

Equation (A5) is an example of a Kesten process (Kesten, 1973), discussed, for example, in Gabaix (1999). Kesten shows that provided that $\frac{1+r-\beta}{1+g} < 1$, (A5) converges to a stationary distribution with a Pareto tail—meaning that the right tail of the distribution, corresponding to $a \geq \bar{a}$ for $\bar{a}$ sufficiently large, can be approximated by $1 - \Gamma a^{-\lambda}$ for some endogenously-determined Pareto shape parameter $\lambda \geq 0$.

To obtain the intuition for why (A5) generates a Pareto tail in the stationary distribution, we consider the following heuristic derivation, which follows Gabaix (1999). Let us focus on the case in which z and ε are iid. Let us also define the counter-cumulative density function of (normalized) wealth in this economy to be $\mathbb{G}(a) \equiv 1 - \Pr[a_{it} \leq a]$. Then

$$\begin{aligned} \Pr[a_{it+1} \geq a] &= \mathbb{E}\left[\mathbf{1}_{\{\gamma a_{it}+z \geq a\}}\right], \\ &= \mathbb{E}\left[\mathbf{1}_{\{a_{it} \geq (a-z)/\gamma\}}\right], \end{aligned}$$

where $\mathbf{1}_{\{\mathcal{P}\}}$ is the indicator function for the event $\mathcal{P}$, we have defined $\gamma \equiv \frac{1+r+\varepsilon-\beta}{1+g}$ for notational convenience, and we have dropped the indices for z and γ since the stochastic laws for these variables do not depend on time and are identical across individuals. Then, by the definition of a stationary distribution $\mathbb{G}$, we have

$$\mathbb{G}(a) = \mathbb{E}\left[\mathbb{G}\left(\frac{a-z}{\gamma}\right)\right].$$

Now let us conjecture a Pareto tail with shape parameter λ, i.e., $\mathbb{G}(a) = \Gamma a^{-\lambda}$ for large a. Then for large a, we have

$$\Gamma a^{-\lambda} = \Gamma \mathbb{E}(a-z)^{-\lambda}\left[\gamma^{\lambda}\right],$$

or

$$1 = \mathbb{E}\left(\frac{a-z}{a}\right)^{-\lambda}\left[\gamma^{\lambda}\right].$$

Since $\mathbb{E}z < \infty$ and has finite variance, we can write $\lim_{a\to\infty} \mathbb{E}\left(\frac{a-z}{a}\right)^{-\lambda} = 1$, which confirms the conjecture and defines λ as the positive solution to

$$\mathbb{E}\left[\gamma^{\lambda}\right] = 1. \tag{A6}$$

This equation also explains why $\mathbb{E}\gamma = \frac{1+r-\beta}{1+g} < 1$ is necessary for convergence to a stationary distribution (as otherwise the wealth distribution would diverge).

Once pinned down, this Pareto shape parameter of the right tail, λ, determines wealth inequality, as well as income inequality, at the top of the distribution. For example, if the entire wealth distribution were Pareto, then the top k's percentile's share of total wealth would be simply: $\left(\frac{k}{100}\right)^{\frac{1-\lambda}{\lambda}}$. This expression makes it clear that a lower λ corresponds to a "thicker tail" of the Pareto distribution and thus to a greater share of aggregate wealth accruing to households in the higher percentiles of the distribution.

The question of interest is whether an increase in $r - g$ (or in $r - g - \beta$) corresponding to a rightward shift in the stochastic distribution of γ will reduce λ, thus leading to greater inequality in the tail of the wealth distribution. Though in general this relationship is ambiguous, in a number of important cases such rightward shifts do reduce λ and increase top inequality as we next show.

Recall that (again ε_{it} and z_{it} being iid), we have

$$a_{it+1} = \gamma_{it} a_{it} + z_{it}.$$

Taking expectations on both sides, using the fact that γ_{it} is iid and that in the stationary distribution $\mathbb{E}a_{it+1} = \mathbb{E}a_{it} = 1$, we have

$$\mathbb{E}\gamma = 1 - \bar{z},$$

where $\bar{z} = \mathbb{E}z_{it} \in (0, 1)$, as noted above. This equation also implies that $\mathbb{E}\gamma \in (0, 1)$.

To determine the relationship between $r - g$ and λ, we consider two special cases.

First suppose that γ (or ε) is log normally distributed. In particular, suppose that $\ln \gamma$ has a normal distribution with mean $\ln(1 - \bar{z}) - \sigma^2/2$ and variance $\sigma^2 > 0$ (so that $\mathbb{E}\gamma = 1 - \bar{z}$). Then we have

$$\mathbb{E}[\gamma^\lambda] = \mathbb{E}[e^{\lambda \ln \gamma}],$$

which is simply the moment generating function of the normally distributed random variable $\ln \gamma$, which can be written as

$$\mathbb{E}[e^{\lambda \ln \gamma}] = e^{\lambda[\ln(1-\bar{z}) - \sigma^2/2] + \lambda^2 \sigma^2/2}.$$

Then the definition of λ, $\mathbb{E}[\gamma^\lambda] = 1$, implies that

$$\lambda[\ln(1 - \bar{z}) - \sigma^2/2] + \lambda^2 \sigma^2/2 = 0,$$

which has two roots, $\lambda = 0$ (which is inadmissible for the stationary distribution), and the relevant one,

$$\lambda = 1 - \frac{\ln(1 - \bar{z})}{\sigma^2/2} > 1.$$

Finally, for small values of $r - g - \beta < 0$, we can write

$$\gamma \approx 1 + r - g - \beta + \varepsilon,$$

and thus from the relationship that $\mathbb{E}\gamma = 1 - \bar{z}$, we have that $\bar{z} = -(r - g - \beta) > 0$, so that

$$\lambda \approx 1 - \frac{\ln(1 + r - g - \beta)}{\sigma^2/2}.$$

It then readily follows that λ is decreasing in $r-g-\beta$, thus implying that higher $r-g$ and lower marginal propensity to consume out of wealth, β, lead to greater top inequality.[14]

Second, a similar relationship can be derived even when γ is not log normally distributed, but only when $\bar{z}$ is small (and we will see why this may not be very attractive in the context of the stationary distribution of wealth). Let us start by taking a first-order Taylor expansion of $\mathbb{E}[\gamma^{\lambda}]=1$ with respect to λ around $\lambda=1$ (which also corresponds to making $\bar{z}$ lie close to zero). In particular, differentiating within the expectation operator, we have

$$\mathbb{E}[\gamma+\gamma\ln\gamma(\lambda-1)]\approx 1,$$

where this approximation requires λ to be close to 1.[15] Then again exploiting the fact that $\mathbb{E}\gamma = 1-\bar{z}$, we have

$$\lambda \approx 1+\frac{\bar{z}}{\mathbb{E}[\gamma\ln\gamma]}>1.$$

(where the fact that $\mathbb{E}[\gamma\ln\gamma]>0$ follows from the fact that $\bar{z}$ is close to zero).[16] This expression clarifies why λ is close to 1 when $\bar{z}$ is close to 0.

Moreover, note that the derivative of $\gamma\ln\gamma$ is $1+\ln\gamma$. For $\bar{z}$ small, $\ln\gamma>-1$ with sufficiently high probability, and thus $\mathbb{E}[\gamma\ln\gamma]$ increases as γ shifts to the right (in the sense of first-order stochastic dominance). Therefore, when λ is close to 1 or equivalently when $\bar{z}$ is close to 0, a higher $r-g-\beta$ increases $\mathbb{E}[\gamma\ln\gamma]$ and reduces the shape parameter λ, raising top inequality. However, it should also be noted that this case is much less relevant for stationary wealth distributions which have Pareto tails much greater than 1.

Benhabib, Bisin and Zhu (2011) extend the result on the Pareto-tail of the wealth distribution to a setup with finitely-lived agents with bequest motives. In this case, the tail of the distribution is in part driven by which individuals have been accumulating for the longest time. They also derive the consumption choices from optimization decisions, consider the equilibrium determination of the interest rate, and confirm the results derived heuristically here. In addition, they show that one type of social mobility—related to the serial correlation of ε, thus making financial returns less correlated for a household over time—tends to make the tail less thick, hence reducing top inequality. These issues are also discussed in Jones (2014).

There are several reasons why these models may not be entirely satisfactory as models of top inequality, however. First, to the extent that very rich individuals have diversified portfolios,

[14] The same conclusion follows without the approximation $\gamma\approx 1+r-g-\beta+\varepsilon$. In this case, we would simply have

$$\lambda = 1-\frac{\ln\left(1+\frac{1+r-\beta}{1+g}\right)}{\sigma^2/2},$$

which yields the same comparative statics.

[15] Formally, we have $\mathbb{E}[\gamma+\gamma\ln\gamma(\lambda-1)+o(\lambda)]=1$.

[16] By noting that $\gamma\ln\gamma$ is a convex function and applying Jensen's inequality, $\mathbb{E}[\gamma\ln\gamma]>\mathbb{E}\gamma\cdot\ln\mathbb{E}\gamma=(1-\bar{z})\ln(1-\bar{z})$. For $\bar{z}$ close enough to zero, $(1-\bar{z})\ln(1-\bar{z})=0$, and thus $\mathbb{E}[\gamma\ln\gamma]>0$.

variability in rates of returns as a driver of the tail of the distribution may not be the most dominant factor. Second, the structure of these models implies that labor income plays no role in the tail of the stationary wealth distribution, but this may be at odds with the importance of wages and salaries and "business income" in the top 1 percent or even top 0.1 percent share of the national income (Piketty and Saez, 2003). Third and relatedly, these models do not have a role for entrepreneurship, which is one of the important aspects of the interplay between labor and capital income (see, for example, Jones and Kim, 2014). Fourth, and most importantly in our opinion, these models do not feature social mobility (except the limited type of social mobility related to the correlation of financial returns considered in Benhabib, Bisin and Zhu, 2011), which appears to be an important determinant of top inequality and its persistence. Finally, in more realistic versions such as Benhabib, Bisin and Zhu (2011) and Jones and Kim (2014), a key determinant of the extent of top inequality turns out to be the age or some other characteristic of the household which determines how long the household has been accumulating. But this is also at odds with the salient patterns of the tail of the income and wealth distribution in the United States, whereby successful entrepreneurs or professionals are more likely to be represented at this tale than individuals or households that have been accumulating capital for a long while.

From $r - g$ to the Implications of Low Growth

A key part of Piketty's argument is that the future will bring even greater inequality because it will be characterized by low economic growth (at least in the developed 'capitalist' economies). This argument relies on two pillars—in addition to the link from $r - g$ to inequality or top inequality as explicated above. The first is that the future will be characterized by low growth. This is not the place to enter into a long debate about the forecasts of future growth rates, but it suffices to note that we do not find forecasts about future growth that do not make any reference to the future of technology, innovation, and the institutions that shape them particularly convincing. Though the demographic trends Piketty emphasizes are well known, their implications for economic growth are much less well understood.

The second important point is that, even if one were to take the link between $r - g$ and inequality on faith, this does not imply that a lower g will translate into a higher $r - g$. As we noted in the text, there are two reasons for this. First, changes in g will impact r from the household side. For example, if consumption decisions are made by optimizing households, then the interest rate is pinned down as $r = \theta g + \rho$, where θ is the inverse of the intertemporal elasticity of substitution. If only some fraction of households optimize and the rest are hand-to-mouth consumers, this equation will still apply because it will be the optimizing consumers who, at the margin, determine the equilibrium interest rate. In cases where $\theta > 1$, $r - g$ would actually decrease with declines in g.

Second, even ignoring the linkage between r and g coming from the household side, changes in g will impact r through their influence on the capital-output ratio (since r is related to the marginal product of capital). This is where Piketty asserts that the elasticity of substitution between capital and labor is very high, ensuring that changes in the capital-labor ratio in the economy do not translate into significant changes in the rate of return to capital and the interest rate. As we noted in the text, however, these strong claims are not backed by the existing evidence. Therefore, we are particularly skeptical of Piketty's conclusion from his theoretical edifice, even with the central role assigned to $r - g$.

These considerations suggest that even if $r - g$ may be a useful statistic in the context of top inequality, it cannot be used either for comparative static type analysis (because it will respond endogenously and depending on technology and institutions to the changes being considered) or even for medium-term forecasting. In addition, the Kaldor-type model presented above highlights another difficulty of reasoning in terms of $r - g$. For this quantity to be constant, we need to specify not only what the saving rate of workers has to be but also how it is changing. In particular, given the saving rate of capitalists and other variables, g is a function of the capitalists's share of national income, ϕ, the saving rate of workers, s_W, and the rate of growth of their labor income. This implies that if $r > g$, then because ϕ is changing, the saving rate and/or the growth rate of labor income of workers must be also implicitly changing.

All of this suggests that r and g must be treated as endogenous variables, and predictions about the future or comparative statics must be conducted in terms of exogenous variables, not in terms of endogenous objects.

Piketty's Second Fundamental Law of Capitalism

The final point we would like to comment on is Piketty's second fundamental law of capitalism, linking the capital-GDP ratio to the saving rate and the growth rate of the economy. Piketty uses this second fundamental law to assert a strong link between the size of the capital stock relative to GDP and the growth rate of the economy, and then on the basis of his forecasts of lower economic growth in the future, reaches the conclusion that the future will bring a pronounced increase in the size of the capital stock relative to GDP in advanced economies. Given a constant interest rate, r, this also implies the continuation of the recent increase in the share of capital in national income. Thus, while the fundamental force of $r - g$ provides an account of a growing top 1 percent share, the second fundamental law of capitalism provides predictions about the future of capital-GDP ratio and the share of national income accruing to capital overall.

In this part of the appendix, we show how something akin to the second fundamental law follows from the Solow growth model, but also that why it is misleading to derive predictions about the evolution of the capital-GDP ratio (or the capital share of national income) from this relationship because it relates these objects to endogenous variables that will all tend to change together in response to shocks or changes in parameters.

Piketty's second fundamental law of capitalism is

$$g = s\frac{Y}{K},$$

where s is the aggregate saving rate. Then, combining this with his first fundamental law (which is just an identity as noted in the text), he obtains that

$$\text{capital share of national income} = \frac{r \times s}{g}.$$

Holding r and s constant, if there is a decline in the growth rate of the economy, g, then capital share of national income will increase. In particular, if the growth rate is halved, then capital's share of national income should double.

Let us start with the steady-state equilibrium of a standard Solow growth model, where there is a constant saving rate, s, and depreciation of capital at the rate δ. Then in this steady-state equilibrium, we have

$$\frac{K}{Y} = \frac{s}{g+\delta}. \tag{A7}$$

To see this, simply note that, assuming a constant saving rate, aggregate saving is

$$sY = I = \dot{K} + \delta K,$$

so that

$$s\frac{Y}{K} = \frac{\dot{K}}{K} + \delta.$$

If we also have $g = \frac{\dot{K}}{K}$, then (A7) follows.

Piketty's version is the special case of this well-known relationship when $\delta = 0$—or when things are specified in "net" units, so that what we have is not national income, but national income net of depreciation, and the saving rate is interpreted as the saving rate above the amount necessary for replenishing depreciated capital. Krusell and Smith (2014) provide a more detailed critique of Piketty's second fundamental law formulated in this way. In particular, as we noted in the text, Piketty's second fundamental law has untenable implications, particularly in the cases where the growth rate of the economy becomes low (and it is these cases on which Piketty bases his conclusions about the implications of low growth on the capital share of national income).

We should also note that the second fundamental law applies when the capital-GDP ratio is constant, and thus $g = \frac{\dot{K}}{K}$ as just noted. Out of steady state (or balance growth path), it is not exactly true. Nevertheless, the relevant conclusion—that there will be an increase in the capital-GDP ratio following a decline in g provided that r and s remain constant—still holds. This follows from the fact that the new steady state following a lower growth rate, say $g' < g$, will involve a higher capital-GDP ratio of

$$\frac{K'}{Y'} = \frac{s}{g' + \delta},$$

and convergence to this new steady state in the baseline Solow model is monotone, so over time the capital-GDP ratio will monotonically increase (though with a small saving rate, the transition can take a long time).

Observe also that because of the depreciation rate, δ, in the denominator, the impact of changes in the growth rate are less than the very large effects Piketty's second fundamental law of capitalism implies (see again Krusell and Smith, 2014).

However, even though we have shown how a version of Piketty's second fundamental law of capitalism follows from the Solow growth model, this does not justify the conclusion that a slowdown in economic growth will automatically increase the capital-GDP ratio or the capital share of national income because, as already noted, almost any change that will reduce the rate of economic growth will also impact the interest rate and the saving rate.

Additional References for the Online Appendix

Aoki, Shuhei and Makoto Nirei (2013) "Pareto Distribution and the Evolution of Top Incomes in the US" University Library of Munich, Paper 47967.

Benhabib, Jess, Alberto Bisin, Shenghao Zhu (2011) "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents" *Econometrica*, 79(1), 123-157.

Gabaix, Xavier (1999) "Zipf's Law for Cities: An Explanation" *Quarterly Journal of Economics*, 114(3), 79-767.

Homburg, Stefan (2014) "Critical Remarks on Piketty's 'Capital in the Twenty-first Century'", Discussion Paper No. 530 (Institute of Public Economics, Leibniz University of Hannover, Germany) and ISSN 0949-9962.

Jones, Charles I. (2014) "The Macroeconomics of Piketty" forthcoming *Journal of Economic Perspectives.*

Jones, Charles I. and Jihee Kim (2014) "A Schumpeterian Model of Top Income Inequality" Stanford mimeo.

Kaldor, Nicholas (1955) "Alternative Theories of Distribution" *Review of Economic Studies*, 23(2), 83-100.

Kesten, Harry (1973) "Random Difference Equations and Renewal Theory for Product to Random Matrices" *Acta Mathematica*, 131, 207-248.

Nirei, Makoto (2009) "Pareto Distributions in Economic Growth Models" Hitotsubashi University, mimeo.

Table 1: Regression coefficients of different proxies of $r-g$. The dependent variable is the top 1 percent share of national income.

	No cross-country variation in r			*OECD data on interest rates*			$r = MPK - \delta$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panel A: Estimates using annual panel					
Estimate of $r-g$ at t	-0.006	-0.018*	-0.018*	-0.066**	-0.038**	-0.040*	0.029	-0.004	-0.011
	(0.012)	(0.010)	(0.011)	(0.027)	(0.017)	(0.021)	(0.033)	(0.009)	(0.008)
Estimate of $r-g$ at $t-1$			0.001			-0.003			0.005
			(0.009)			(0.015)			(0.014)
Estimate of $r-g$ at $t-2$			0.005			0.010			-0.012
			(0.008)			(0.019)			(0.008)
Estimate of $r-g$ at $t-3$			-0.002			-0.012			0.014*
			(0.008)			(0.024)			(0.008)
Estimate of $r-g$ at $t-4$			-0.005			-0.005			0.006
			(0.007)			(0.013)			(0.010)
Joint significance of lags [p-value]			4.55 [0.47]			7.47 [0.19]			12.40 [0.03]
Long-run effect [p-value estimate> 0]		-0.16 [0.13]	-0.18 [0.15]		-0.39 [0.29]	-0.47 [0.34]		-0.04 [0.68]	0.03 [0.89]
Persistence of top 1 percent share [p-value estimate< 1]		0.89 [0.00]	0.89 [0.00]		0.90 [0.31]	0.89 [0.30]		0.90 [0.11]	0.92 [0.18]
Observations	1646	1233	1226	627	520	470	1162	905	860
Countries	27	27	27	19	18	18	28	26	26
				Panel B: Estimates using 10-year and 20-year panels (columns 3,6,9)					
Average $r-g$	0.055	-0.036	-0.252	-0.114	-0.121	-0.110	0.069	0.148	0.238
	(0.110)	(0.118)	(0.269)	(0.138)	(0.132)	(0.320)	(0.118)	(0.100)	(0.164)
Long-run effect [p-value estimate> 0]		-0.05 [0.76]			-0.25 [0.44]			0.29 [0.22]	
Persistence of top 1 percent share [p-value estimate< 1]		0.32 [0.00]			0.52 [0.02]			0.48 [0.00]	
Observations	213	181	106	82	80	43	135	124	61
Countries	27	25	24	18	18	17	27	25	22

Notes-: The table presents estimates of different proxies of $r-g$ on the top 1 percent share of national income. The dependent variable is available from 1871 onwards for the countries covered in the World Top Incomes Database. We use different proxies of $r-g$: Columns 1 to 3 use growth rates from Madisson, and assume no variation in real interest rates across countries. These data are available from 1870 onwards. Columns 4 to 6 use real interest rates computed by subtracting realized inflation from nominal yields on long-term government bonds, and growth rates from the Penn World Tables. These data are only available since 1955 for OECD countries. Columns 7 to 9 use $r = MPK - \delta$, constructed as explained in the text using data from the Penn World Tables, and growth rates from the Penn World Tables. These data are available for 1950 onwards. Panel A uses an unbalanced yearly panel. Columns 2,5 and 8 add five lags of the dependent variable and report the estimated persistence of the top 1 percent share of national income and the estimated long run effect of $r-g$ on the dependent variable. Columns 3,6 and 9 add four lags of $r-g$ on the right-hand side, and also report the long-run effect of a permanent increase of 1% in $r-g$ and a test for the joint significance of these lags (with its corresponding χ^2 statistic and p-value). Panel B uses an unbalanced panel with observations every 10 years or 20 years (columns 3,6,9). Columns 1,2,4,5,7 and 8 present estimates from a regression of the top 1 percent share of national income at the end of each decade in the sample (that is, 1880, 1890, ..., 2010, depending on data availability) on the average $r-g$ during the decade. Columns 2,5, and 8 add one lag of the dependent variable on the right-hand side. Finally, columns 3,6 and 9, present estimates from a regression of the top 1 percent share of national income at the end of each 20-year period in the sample (that is, 1890, 1910, ..., 2010, depending on data availability) on the average $r-g$ during the period. All specifications include a full set of country and year fixed effects. Standard errors allowing for arbitrary heteroscedasticity and serial correlation of residuals at the country level are computed using the pairs-cluster bootstrap procedure proposed by Cameron, Gelbach and Miller (2008) and are reported in parentheses.

Table A1: Regression coefficients of different proxies of $r - g$. The dependent variable is the top 1 percent share of national income. Traditional standard errors assuming homoscedasticity and no serial correlation.

	No cross-country variation in r			*OECD data on interest rates*			$r = MPK - \delta$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Estimates using annual panel									
Estimate of $r - g$ at t	-0.006	-0.018***	-0.018***	-0.066***	-0.038***	-0.040**	0.029*	-0.004	-0.011
	(0.011)	(0.005)	(0.005)	(0.022)	(0.013)	(0.017)	(0.016)	(0.008)	(0.010)
Estimate of $r - g$ at $t - 1$			0.001			-0.003			0.005
			(0.006)			(0.019)			(0.010)
Estimate of $r - g$ at $t - 2$			0.005			0.010			-0.012
			(0.006)			(0.019)			(0.010)
Estimate of $r - g$ at $t - 3$			-0.002			-0.012			0.014
			(0.006)			(0.019)			(0.010)
Estimate of $r - g$ at $t - 4$			-0.005			-0.005			0.006
			(0.006)			(0.017)			(0.009)
Joint significance of lags [p-value]			2.65 [0.02]			1.53 [0.18]			1.01 [0.41]
Long-run effect [p-value estimate> 0]		-0.16 [0.00]	-0.18 [0.05]		-0.39 [0.03]	-0.47 [0.06]		-0.04 [0.67]	0.03 [0.89]
Persistence of top 1 percent share [p-value estimate< 1]		0.89 [0.00]	0.89 [0.00]		0.90 [0.00]	0.89 [0.00]		0.90 [0.00]	0.92 [0.00]
Observations	1646	1233	1226	627	520	470	1162	905	860
Countries	27	27	27	19	18	18	28	26	26
Years per country	61.0	45.7	45.4	33.0	28.9	26.1	41.5	34.8	33.1
Panel B: Estimates using 10-year and 20-year panels (columns 3,6,9)									
Average $r - g$	0.055	-0.036	-0.252	-0.114	-0.121	-0.110	0.069	0.148*	0.238
	(0.095)	(0.098)	(0.228)	(0.132)	(0.118)	(0.247)	(0.091)	(0.088)	(0.172)
Long-run effect [p-value estimate> 0]		-0.05 [0.72]			-0.25 [0.32]			0.29 [0.11]	
Persistence of top 1 percent share [p-value estimate< 1]		0.32 [0.00]			0.52 [0.00]			0.48 [0.00]	
Observations	213	181	106	82	80	43	135	124	61
Countries	27	25	24	18	18	17	27	25	22
Years per country	7.9	7.2	4.4	4.6	4.4	2.5	5.0	5.0	2.8

Notes-: The table presents estimates of different proxies of $r-g$ on the top 1 percent share of national income. The dependent variable is available from 1871 onwards for the countries covered in the World Top Incomes Database. We use different proxies of $r - g$: Columns 1 to 3 use growth rates from Madisson, and assume no variation in real interest rates across countries. These data are available from 1870 onwards. Columns 4 to 6 use real interest rates computed by subtracting realized inflation from nominal yields on long-term government bonds, and growth rates from the Penn World Tables. These data are only available since 1955 for OECD countries. Columns 7 to 9 use $r = MPK - \delta$, constructed as explained in the text using data from the Penn World Tables, and growth rates from the Penn World Tables. These data are available for 1950 onwards. Panel A uses an unbalanced yearly panel. Columns 2,5 and 8 add five lags of the dependent variable and report the estimated persistence of the top 1 percent share of national income and the estimated long run effect of $r - g$ on the dependent variable. Columns 3,6 and 9 add four lags of $r - g$ on the right-hand side, and also report the long-run effect of a permanent increase of 1% in $r - g$ and a test for the joint significance of these lags (with its corresponding χ^2 statistic and p-value). Panel B uses an unbalanced panel with observations every 10 years or 20 years (columns 3,6,9). Columns 1,2,4,5,7 and 8 present estimates from a regression of the top 1 percent share of national income at the end of each decade in the sample (that is, 1880, 1890, . . . , 2010, depending on data availability) on the average $r - g$ during the decade. Columns 2,5, and 8 add one lag of the dependent variable on the right-hand side. Finally, columns 3,6 and 9, present estimates from a regression of the top 1 percent share of national income at the end of each 20-year period in the sample (that is, 1890, 1910, . . . , 2010, depending on data availability) on the average $r - g$ during the period. All specifications include a full set of country and year fixed effects. Traditional standard errors, imposing homoscedasticity and no residual auto correlation, are reported in parentheses.

Table A2: Regression coefficients of different proxies of $r-g$ controlling for GDP per capita, population growth and country trends.

	No variation in r		*OECD interest rates*		$r = MPK - \delta$	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Baseline			
Estimate of $r-g$ at t	-0.006	-0.018*	-0.066**	-0.038**	0.029	-0.004
	(0.012)	(0.010)	(0.027)	(0.017)	(0.033)	(0.009)
Long-run effect [p-value estimate> 0]		-0.16 [0.13]		-0.39 [0.29]		-0.04 [0.68]
Persistence of top 1 percent share [p-value estimate< 1]		0.89 [0.00]		0.90 [0.31]		0.90 [0.11]
Observations	1646	1233	627	520	1162	905
Countries	27	27	19	18	28	26
			Panel B: log of GDP per capita			
Estimate of $r-g$ at t	-0.006	-0.018*	-0.035	-0.039**	0.032	-0.006
	(0.011)	(0.010)	(0.028)	(0.017)	(0.031)	(0.009)
log GDP per capita at t	-0.169	0.022	3.270	-0.096	0.145	-0.199
	(0.767)	(0.166)	(2.149)	(0.809)	(1.152)	(0.281)
Long-run effect [p-value estimate> 0]		-0.16 [0.14]		-0.41 [0.36]		-0.06 [0.55]
Persistence of top 1 percent share [p-value estimate< 1]		0.89 [0.00]		0.91 [0.35]		0.90 [0.14]
Observations	1646	1233	620	514	1151	898
Countries	27	27	19	18	28	26
			Panel C: Population growth			
Estimate of $r-g$ at t	0.004	-0.017*	-0.039	-0.034*	0.030	-0.006
	(0.013)	(0.009)	(0.027)	(0.018)	(0.031)	(0.008)
Population growth at t	0.255	0.033	0.544	0.117	0.140	-0.055
	(0.225)	(0.060)	(0.464)	(0.139)	(0.310)	(0.067)
Long-run effect [p-value estimate> 0]		-0.15 [0.11]		-0.37 [0.34]		-0.05 [0.51]
Persistence of top 1 percent share [p-value estimate< 1]		0.89 [0.00]		0.91 [0.30]		0.90 [0.10]
Observations	1646	1233	608	503	1134	885
Countries	27	27	19	18	27	26
			Panel D: Country trends			
Estimate of $r-g$ at t	-0.002	-0.018*	-0.022	-0.024	0.015	-0.006
	(0.010)	(0.011)	(0.015)	(0.017)	(0.016)	(0.009)
Long-run effect [p-value estimate> 0]		-0.10 [0.15]		-0.06 [0.07]		-0.02 [0.52]
Persistence of top 1 percent share [p-value estimate< 1]		0.82 [0.00]		0.62 [0.00]		0.70 [0.00]
Observations	1646	1233	627	520	1162	905
Countries	27	27	19	18	28	26

Notes-: The table presents estimates of different proxies of $r-g$ on the top 1 percent share of national income. The dependent variable is available from 1871 onwards for the countries covered in the World Top Incomes Database. We use different proxies of $r-g$: Columns 1 and 2 use growth rates from Madisson, and assume no variation in real interest rates across countries. These data are available from 1870 onwards. Columns 3 and 4 use real interest rates computed by subtracting realized inflation from nominal yields on long-term government bonds, and growth rates from the Penn World Tables. These data are only available since 1955 for OECD countries. Columns 5 and 6 use $r = MPK - \delta$, constructed as explained in the text using data from the Penn World Tables, and growth rates from the Penn World Tables. These data are available for 1950 onwards. Columns 2,4 and 6 add five lags of the dependent variable and report the estimated persistence of the top 1 percent share of national income and the estimated long run effect of $r-g$ on the dependent variable. Panel A presents the baseline estimates. Panel B adds the log of GDP per capita as a control. Panel C adds population growth as a control. Finally, Panel D adds country-specific trends as controls. All specifications include a full set of country and year fixed effects. Standard errors allowing for arbitrary heteroscedasticity and serial correlation of residuals at the country level are computed using the pairs-cluster bootstrap procedure proposed by Cameron, Gelbach and Miller (2008) and are reported in parentheses.

Table A3: Regression coefficients of different proxies of $r-g$. The dependent variable is the top 5 percent share of national income.

	No cross-country variation in r			*OECD data on interest rates*			$r = MPK - \delta$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Estimates using annual panel									
Estimate of $r-g$ at t	-0.002	0.010	0.007	-0.109**	-0.039	-0.046	0.056	0.006	-0.006
	(0.034)	(0.022)	(0.022)	(0.049)	(0.028)	(0.033)	(0.068)	(0.022)	(0.027)
Estimate of $r-g$ at $t-1$			-0.001			0.008			0.005
			(0.018)			(0.021)			(0.021)
Estimate of $r-g$ at $t-2$			0.035***			0.007			-0.007
			(0.013)			(0.025)			(0.016)
Estimate of $r-g$ at $t-3$			-0.006			0.010			0.020
			(0.013)			(0.034)			(0.012)
Estimate of $r-g$ at $t-4$			-0.008			-0.001			0.011
			(0.013)			(0.019)			(0.015)
Joint significance of lags [p-value]			14.81 [0.01]			3.63 [0.60]			4.63 [0.46]
Long-run effect [p-value estimate> 0]		0.12 [0.67]	0.34 [0.37]		-0.48 [0.30]	-0.25 [0.66]		0.08 [0.80]	0.34 [0.41]
Persistence of top 1 percent share [p-value estimate< 1]		0.92 [0.01]	0.92 [0.01]		0.92 [0.28]	0.91 [0.27]		0.93 [0.12]	0.93 [0.13]
Observations	1307	988	988	590	489	440	985	786	749
Countries	24	21	21	18	17	17	24	20	20
Panel B: Estimates using 10-year and 20-year panels (columns 3,6,9)									
Average $r-g$	-0.019	-0.147	-0.602	-0.151	-0.075	-0.043	0.102	0.252**	0.323
	(0.207)	(0.207)	(0.514)	(0.224)	(0.217)	(0.500)	(0.199)	(0.128)	(0.256)
Long-run effect [p-value estimate> 0]		-0.24 [0.49]			-0.19 [0.73]			0.50 [0.14]	
Persistence of top 1 percent share [p-value estimate< 1]		0.39 [0.00]			0.60 [0.12]			0.50 [0.00]	
Observations	171	143	86	78	76	41	114	105	55
Countries	22	21	20	17	17	16	22	21	20

Notes-: The table presents estimates of different proxies of $r-g$ on the top 5 percent share of national income. The dependent variable is available from 1871 onwards for the countries covered in the World Top Incomes Database. We use different proxies of $r-g$: Columns 1 to 3 use growth rates from Madisson, and assume no variation in real interest rates across countries. These data are available from 1870 onwards. Columns 4 to 6 use real interest rates computed by subtracting realized inflation from nominal yields on long-term government bonds, and growth rates from the Penn World Tables. These data are only available since 1955 for OECD countries. Columns 7 to 9 use $r = MPK - \delta$, constructed as explained in the text using data from the Penn World Tables, and growth rates from the Penn World Tables. These data are available for 1950 onwards. Panel A uses an unbalanced yearly panel. Columns 2,5 and 8 add five lags of the dependent variable and report the estimated persistence of the top 5 percent share of national income and the estimated long run effect of $r-g$ on the dependent variable. Columns 3,6 and 9 add four lags of $r-g$ on the right-hand side, and also report the long-run effect of a permanent increase of 1% in $r-g$ and a test for the joint significance of these lags (with its corresponding χ^2 statistic and p-value). Panel B uses an unbalanced panel with observations every 10 years or 20 years (columns 3,6,9). Columns 1,2,4,5,7 and 8 present estimates from a regression of the top 5 percent share of national income at the end of each decade in the sample (that is, 1880, 1890, ..., 2010, depending on data availability) on the average $r-g$ during the decade. Columns 2,5, and 8 add one lag of the dependent variable on the right-hand side. Finally, columns 3,6 and 9, present estimates from a regression of the top 5 percent share of national income at the end of each 20-year period in the sample (that is, 1890, 1910, ..., 2010, depending on data availability) on the average $r-g$ during the period. All specifications include a full set of country and year fixed effects. Standard errors allowing for arbitrary heteroscedasticity and serial correlation of residuals at the country level are computed using the pairs-cluster bootstrap procedure proposed by Cameron, Gelbach and Miller (2008) and are reported in parentheses.

Table A4: Regression coefficients of different proxies of $r-g$. The dependent variable is the top 1 percent share of national income. Sample restricted to OECD countries since 1950.

	No cross-country variation in r			*OECD data on interest rates*			$r = MPK - \delta$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panel A: Estimates using annual panel					
Estimate of $r-g$ at t	-0.127***	-0.057***	-0.057**	-0.066**	-0.038**	-0.040*	0.074	-0.020	-0.020
	(0.046)	(0.022)	(0.023)	(0.027)	(0.017)	(0.021)	(0.080)	(0.015)	(0.020)
Estimate of $r-g$ at $t-1$			-0.001			-0.003			-0.001
			(0.030)			(0.015)			(0.039)
Estimate of $r-g$ at $t-2$			-0.014			0.010			-0.009
			(0.020)			(0.019)			(0.022)
Estimate of $r-g$ at $t-3$			-0.014			-0.012			0.008
			(0.023)			(0.024)			(0.019)
Estimate of $r-g$ at $t-4$			-0.025			-0.005			-0.020
			(0.021)			(0.013)			(0.030)
Joint significance of lags [p-value]			13.47 [0.02]			7.47 [0.19]			3.34 [0.65]
Long-run effect [p-value estimate> 0]		-0.61 [0.31]	-1.02 [0.32]		-0.39 [0.29]	-0.47 [0.34]		-0.24 [0.53]	-0.47 [0.63]
Persistence of top 1 percent share [p-value estimate< 1]		0.91 [0.34]	0.89 [0.33]		0.90 [0.31]	0.89 [0.30]		0.91 [0.39]	0.91 [0.44]
Observations	627	520	470	627	520	470	627	520	470
Countries	19	18	18	19	18	18	19	18	18
				Panel B: Estimates using 10-year and 20-year panels (columns 3,6,9)					
Average $r-g$	-0.671***	-0.631***	-1.146*	-0.114	-0.121	-0.110	0.052	0.017	0.132
	(0.256)	(0.208)	(0.599)	(0.138)	(0.132)	(0.320)	(0.208)	(0.157)	(0.279)
Long-run effect [p-value estimate> 0]		-1.30 [0.11]			-0.25 [0.44]			0.04 [0.92]	
Persistence of top 1 percent share [p-value estimate< 1]		0.51 [0.02]			0.52 [0.02]			0.53 [0.03]	
Observations	82	80	43	82	80	43	82	80	43
Countries	18	18	17	18	18	17	18	18	17

Notes-: The table presents estimates of different proxies of $r-g$ on the top 1 percent share of national income. We restrict our sample to OECD countries for which interest rates data is available from 1955 onwards. The countries in our sample include Australia, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, South Africa, Spain, Sweden, Switzerland, United Kingdom and United States. We use different proxies of $r-g$: Columns 1 to 3 use growth rates from the Penn World Tables, and assume no variation in real interest rates across countries. Columns 4 to 6 use real interest rates computed by subtracting realized inflation from nominal yields on long-term government bonds, and growth rates from the Penn World Tables. Columns 7 to 9 use $r = MPK - \delta$, constructed as explained in the text using data from the Penn World Tables, and growth rates from the Penn World Tables. Panel A uses an unbalanced yearly panel. Columns 2,5 and 8 add five lags of the dependent variable and report the estimated persistence of the top 1 percent share of national income and the estimated long run effect of $r-g$ on the dependent variable. Columns 3,6 and 9 add four lags of $r-g$ on the right-hand side, and also report the long-run effect of a permanent increase of 1% in $r-g$ and a test for the joint significance of these lags (with its corresponding χ^2 statistic and p-value). Panel B uses an unbalanced panel with observations every 10 years or 20 years (columns 3,6,9). Columns 1,2,4,5,7 and 8 present estimates from a regression of the top 1 percent share of national income at the end of each decade in the sample (that is, 1880, 1890, ..., 2010, depending on data availability) on the average $r-g$ during the decade. Columns 2,5, and 8 add one lag of the dependent variable on the right-hand side. Finally, columns 3,6 and 9, present estimates from a regression of the top 1 percent share of national income at the end of each 20-year period in the sample (that is, 1970, 1990, ..., 2010) on the average $r-g$ during the period. All specifications include a full set of country and year fixed effects. Standard errors allowing for arbitrary heteroscedasticity and serial correlation of residuals at the country level are computed using the pairs-cluster bootstrap procedure proposed by Cameron, Gelbach and Miller (2008) and are reported in parentheses.

Table A5: Regression coefficients of different proxies of $r - g$. The dependent variable is the capital share of national income.

	Dep. var: capital share from Penn World Tables						*Dep. var: capital share from Karabarbounis and Neiman (2013)*					
	No cross-country variation in r			*OECD data on interest rates*			*No cross-country variation in r*			*OECD data on interest rates*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Estimates using annual panel												
Estimate of $r - g$ at t	-0.045	-0.009	-0.008	-0.240***	-0.053**	-0.072***	-0.203***	-0.052**	-0.075***	-0.100	-0.007	-0.033
	(0.033)	(0.011)	(0.011)	(0.088)	(0.026)	(0.025)	(0.057)	(0.025)	(0.025)	(0.093)	(0.031)	(0.029)
Estimate of $r - g$ at $t - 1$			-0.004			0.046*			0.088***			0.059*
			(0.011)			(0.025)			(0.021)			(0.033)
Estimate of $r - g$ at $t - 2$			-0.005			0.063*			0.023			0.062**
			(0.007)			(0.036)			(0.025)			(0.025)
Estimate of $r - g$ at $t - 3$			-0.006			-0.002			-0.018			-0.013
			(0.007)			(0.033)			(0.022)			(0.031)
Estimate of $r - g$ at $t - 4$			-0.006			0.048*			0.014			0.031
			(0.007)			(0.026)			(0.019)			(0.023)
Joint significance of lags [p-value]			. [0.81]			. [0.00]			. [0.00]			. [0.00]
Long-run effect [p-value estimate> 0]		-0.05 [0.41]	-0.15 [0.19]		-0.28 [0.03]	0.46 [0.13]		-0.20 [0.02]	0.14 [0.51]		-0.03 [0.82]	0.45 [0.06]
Persistence of capital share [p-value estimate< 1]		0.81 [0.00]	0.81 [0.00]		0.81 [0.00]	0.82 [0.00]		0.74 [0.00]	0.77 [0.00]		0.76 [0.00]	0.76 [0.00]
Observations	2687	2619	2611	412	412	397	1735	1239	1239	495	430	406
Countries	123	123	123	19	19	19	99	92	92	19	19	19
Panel B: Estimates using 10-year and 20-year panels (columns 3,6,9,12)												
Average $r - g$	-0.137	-0.136	-0.258	0.207	0.187	0.560*	-0.026	-0.534	-0.269	0.254	0.086	-0.046
	(0.124)	(0.122)	(0.307)	(0.185)	(0.196)	(0.331)	(0.156)	(0.339)	(0.742)	(0.275)	(0.385)	(0.374)
Long-run effect [p-value estimate> 0]		-0.16 [0.27]			0.18 [0.36]			-0.53 [0.13]			0.06 [0.83]	
Persistence of capital share [p-value estimate< 1]		0.13 [0.00]			-0.06 [0.00]			-0.01 [0.00]			-0.37 [0.00]	
Observations	350	350	208	55	55	34	151	59	40	56	39	26
Countries	123	123	104	19	19	17	57	22	20	18	14	13

Notes-: The table presents estimates of different proxies of $r - g$ on the capital share of national income. The dependent variable is the capital share of national income. In columns 1 to 6, we use data from the Penn World Tables to compute the capital share for 1990 onwards. In columns 7 to 12, we use the capital share data from Karabarbounis and Neiman (2013). We use different proxies of $r - g$: Columns 1 to 3 and 7 to 9 use growth rates from Madisson, and assume no variation in real interest rates across countries. These data are available from 1870 onwards for most of the countries in the sample. Columns 4 to 6 and 10 to 12 use real interest rates computed by subtracting realized inflation from nominal yields on long-term government bonds, and growth rates from the Penn World Tables. These data are only available since 1955 for OECD countries. Panel A uses an unbalanced yearly panel. Columns 2,5,8 and 11 add five lags of the dependent variable and report the estimated persistence of the capital share of national income and the estimated long run effect of $r - g$ on the dependent variable. Columns 3,6,9 and 12 add four lags of $r - g$ on the right-hand side, and also report the long-run effect of a permanent increase of 1% in $r - g$ and a test for the joint significance of these lags (with its corresponding χ^2 statistic and p-value). Panel B uses an unbalanced panel with observations every 10 years or 20 years (columns 3,6,9,12). Columns 1,2,4,5,7,8,10 and 11 present estimates from a regression of the capital share of national income at the end of each decade in the sample (that is, 1980, 1990, . . . , 2010, depending on data availability) on the average $r - g$ during the decade. Columns 2,5,8 and 11 add one lag of the dependent variable on the right-hand side. Finally, columns 3,6,9 and 12, present estimates from a regression of the capital share of national income at the end of each 20-year period in the sample (that is, 1990 and 2010) on the average $r - g$ during the period. All specifications include a full set of country and year fixed effects. Standard errors allowing for arbitrary heteroscedasticity and serial correlation of residuals at the country level are computed using the pairs-cluster bootstrap procedure proposed by Cameron, Gelbach and Miller (2008) and are reported in parentheses.