Wealth inequality in the long run:  
A Schumpeterian growth perspective  

Jakob B. Madsen,∗ Antonio Minniti, Francesco Venturini  

May 12, 2020  

Abstract  
This paper extends the analysis of the wealth-income ratio based on the neoclassical model in a Schumpeterian growth framework in which savings are channelled to both tangible and intangible capital investment. Using historical data for 21 OECD countries over the period 1860-2015, we find that the wealth-income ratio and, hence, wealth inequality, is negatively related to the rate of economic growth and positively related to the rates of investment in intangible and tangible assets, as predicted by the theory. Accounting for the innovation-induced counteracting growth-effect on the wealth-income ratio, we show that the net effect of investment in intangibles on wealth inequality is positive. Our estimates suggest that intangibles have been a contributing factor in wealth inequality since 1860 and that the marked increase in the investment in intangible assets in the post-WWII period has been a significant driver of wealth inequality since the 1970s.  

Keywords: Wealth-Income Ratio, Intangibles, Tangible Capital, Schumpeterian Growth  

JEL Classification: D30, E10, E20, O30, O40  

Short title: Wealth inequality in long run

*Corresponding Author: Jakob B. Madsen, Department of Economics, University of Western Australia, Perth, WA 6009, Australia & Centre for Applied Macroeconomic Analysis, ANU (jakob.madsen@uwa.edu.au). Helpful comments and suggestions from Matteo Cervellati, Vincenzo Denicolò, Jan Ditzen, Richard Frensch, Stephen Hall, Praveen Kujal, Nick Oulton, Rebecca Riley, James Rockey, Milica Uvalic, Michela Vecchi, Piercarlo Zanchettin, participants at seminars at University of Aix-Marseille, Middlesex Business School, University of Leicester, University of Perugia, University of Bologna, National Institute of Economic and Social Research, London, Hiroshima University, the Asset Conference, Thessaloniki, the VII Workshop on Institutions, Individual Behavior and Economic Outcomes, 2016, Alghero, the Workshop on Economic Growth and Macroeconomics, Sophia-Antipolis, the 14th biannual EACES Conference, Regensburg, the GGDC 25th-anniversary Conference, Groningen and, particularly, the editor (Morten Ravn) and two referees are gratefully acknowledged. Jakob B. Madsen gratefully acknowledges financial support from the Australian Research Council (grant # DP170100339). Antonio Minniti gratefully acknowledges financial support from the Italian Ministry for the University and Scientific Research (PRIN 2015 - 20157NH5TP_004).
1 Introduction

The macroeconomic research on inequality increasingly emphasizes physical and human capital as the driving forces of factor income shares and the wealth-income ratio in industrialized countries (see, e.g., Karabarbounis and Neiman 2014; Piketty and Zucman 2014; Grossman et al. 2017; Acemoglu and Restrepo 2018). Focusing on the trends in the labour share since the 1970s, Karabarbounis and Neiman (2014) argue that declining relative prices of investment goods has induced firms to use more capital at the expense of labour; thus reducing the labour share. Piketty and Zucman (2014) suggest that a stable savings rate combined with reduced income growth is behind the recent increase in the wealth-income ratio, $W/Y$. Using a neoclassical growth setting, Grossman et al. (2017) argue that the productivity slowdown since the 1970s has generated a deceleration in educational attainment and caused an increase in capital’s share of income in the US. Acemoglu and Restrepo (2018) show that the recent decline in the labour share in the US could be a consequence of automation of tasks previously performed by labour.

We contribute to this line of research by showing that intangibles have been a contributing factor in wealth inequality since 1860 and that the marked increase in investment in intangible assets has been a significant driver of the increasing inequality since the 1970s. To this end, we extend the analysis of the $W-Y$ ratio to allow for innovations, noting that the extant literature based on the neoclassical growth model has predominantly focussed on tangible capital (Madsen 2019). The extension is relevant for modern economies in which wealth holders increasingly derive their income from intangible assets (e.g., R&D, copyrights, skills, brands, organizational know-how) rather than from the accumulation of physical capital (Corrado and Hulten 2010). For the US, for example, in 2010 intangible assets accounted for 50% of total investment, and the markedly increasing share of intangibles in the OECD suggests that the 50% mark will soon be reached in other OECD countries (Corrado et al. 2013).

In the paper we provide a Schumpeterian interpretation of the evolution of the $W-Y$ ratio by deriving a steady-state condition for the $W-Y$ ratio within an innovation-driven growth model. More specifically, we extend the canonical quality-ladder model developed by Grossman and Helpman (1991) to include physical capital accumulation and two classes of agents, namely capitalists – who derive their income from accumulated tangible and intangible capital, and workers – who derive their income from labour only. In

\[ W/Y = \frac{s}{(g + \delta)} \]

Piketty and Zucman (2014), Piketty (2015b) and Madsen et al. (2018b) use this condition to analyse the evolution of the wealth-income ratio as a proxy for wealth inequality for advanced economies. Note that Piketty (2015b) p. 49) incorporates “immaterial capital” into his definition of capital; thus treating tangibles and intangibles indistinguishably.

Haskel and Westlake (2017) show that intangible capital has dominated the transition towards the weightless economy, playing an important role in explaining some of the big economic changes of the last decade. See Peretto (2015, 2017) for a theoretical exposition.
this model, technological progress resulting from costly and deliberative research aimed at developing higher quality products is the major engine of growth. The model predicts that the $W$-$Y$ ratio, $\beta$, is negatively related to the rate of economic growth, $g$, and positively related to investment in intangibles, $s_I$, and tangibles, $s_T$, both expressed as a percentage of GDP.

Drawing on this growth-theoretic framework, we carry out a long-run econometric analysis for 21 OECD countries using annual data over the period 1860-2015 to test the predictions of the model, focusing mainly on the effects of intangible investment on the $W$-$Y$ ratio. We show that $\beta$ is positively related to $s_I$ and $s_T$ and negatively related to $g$ in the long run. We base our analysis on dynamic techniques of panel regression and evaluate the robustness of our main results to variation in estimation period, to the inclusion of controls, to various measurement issues exploiting low frequency (five-year interval) data and to simultaneity feedbacks using an instrumental variables (IV) regression approach in which the exogenous variation from natural disasters is exploited.

Furthermore, we undertake an in-depth analysis of the long-run evolution of the $W$-$Y$ ratio by investigating the multiple channels through which investment in intangibles affects wealth inequality. This issue is of fundamental importance because, in the neoclassical growth framework, an increase in the savings ratio is unambiguously associated with a higher $W$-$Y$ ratio in steady state. In the Schumpeterian model, the increasing investment in intangibles has two counterbalancing effects on the $W$-$Y$ ratio. On the one hand, it raises the market value of innovative firms, leading to an increase in the wealth-income ratio (wealth channel). On the other hand, a higher rate of innovation, induced by investment in intangibles, increases the rate of income growth, which ultimately reduces the $W$-$Y$ ratio (growth channel). To track both mechanisms of transmission from innovation to the $W$-$Y$ ratio, we estimate a simultaneous system of equations based on the 3SLS estimator, which explicitly accounts for both the wealth and growth effects. These estimates are then used to evaluate the overall (net) effect of the investment in intangibles on the $W$-$Y$ ratio. Allowing for the innovation-induced counteracting growth-effect on the $W$-$Y$ ratio, our estimates show that the net effect of investment in intangibles on wealth inequality is positive. The 12-fold increase in the intangible investment ratio in the post-WWII period has contributed a 53% increase in the $W$-$Y$ ratio. Conversely, the decline in the net non-residential investment ratio since 1964 (1980) has resulted in a 26% (19%) decline in the $W$-$Y$ ratio, suggesting that tangible capital has not been a source of the increasing wealth inequality over the past four decades.

The paper makes two principal contributions to the literature. First, our work gives a theoretical foundation to the evolution of the $W$-$Y$ ratio for modern economies in which the share of intangibles in total capital is currently accelerating. This extension is important since innovation and intangible capital investments are the main drivers of growth in advanced economies (Madsen 2008), while traditional (tangible) capital investment was the primary driver of economic progress in the 19th century (Galor and Moav 2004). Second, we construct a large macro dataset for 21 OECD countries spanning the period 1860-2015. A significant contribution is the construction of historical R&D expenditure, and related

\footnote{See [Acemoglu and Robinson 2015], [Ray 2015] and [Krusell and Smith 2015] for a discussion of the endogeneity problem concerning Piketty’s theoretical analysis.}
deflators, for which official data are first available for the post-WWII period. As detailed in the data section and the online Appendix, the historical R&D data are derived from higher education sector R&D, government and private expenditure on tertiary education and gross enrolment rates in tertiary education, and have been compiled from several different data sources and documents going back in time.

The long historical data give several advantages over samples spanning only a few decades. First, the data trace the determinants of the \( W-Y \) ratio from the beginning of the Second Industrial Revolution with high wealth inequality, through the industrial era with low inequality, and the transition into the post-industrial regime in which wealth inequality is gradually converging to the level that prevailed before the end of WWI (Piketty and Zucman 2014; Roine and Waldenström 2015). Second, since the adjustment towards steady state is a slow process, estimates covering only a few decades may be biased because the parameters are driven by a transitional path and, thus, do not capture the long-run structural relationship - a problem that, to a large extent, is overcome with long data. Third, long panel data give crucial econometric advantages: the panel dynamic estimator becomes more consistent as the sample grows and IV estimates are substantially more consistent in large than in small samples (Bekker 1994; Hahn and Hausman 2005; Powell 2017).

In addition to the contributions of Karabarbounis and Neiman (2014), Piketty and Zucman (2014), Grossman et al. (2017), our paper relates to a growing literature on the evolution of factor income shares and their determinants (see, e.g., Elsby et al. 2013; Bridgman 2018; Rogné 2015; Lawrence 2015; Barkai forthcoming; Kehrig and Vincent 2017; Koh et al. 2018). Finally, our paper is also related to a recent strand of the literature that uses a Schumpeterian growth approach to analyse the relationship between innovation and top income inequality, and between foreign technological competition, innovation, and the \( W-Y \) ratio. Jones and Kim (2018) and Aghion et al. (2019) investigate the macroeconomic implications of the increasing share of income going to top income earners. Cozzi and Impullitti (2016) show, by means of calibrations, that the shrinking technological lead of the US has been responsible for a significant increase in the \( W-Y \) ratio in the US.

The paper is structured as follows. Section 2 lays out the growth set-up which provides an organizing framework for the empirical analysis. Section 3 describes the data and presents the econometric model and the estimation procedure. Regression results and robustness checks are presented in Section 4. Section 5 investigates the net effects of investment in R&D and other intangibles on wealth inequality. Finally, Section 6 concludes.

---

\(^{5}\)Schumpeterian growth models are characterized by endogenous factor income shares and, therefore, constitute an appropriate framework for studying the relationship between income distribution and economic growth (see, e.g., Peretto 2015; Irmen and Tabakov forthcoming; Growiec et al. 2018). Moreover, to the extent to which these models incorporate different ownership shares of firms across classes of agents, they can be used to study the link between personal distribution of income and wealth and economic growth (see, e.g., Chou and Talmain 1996; Zweimüller 2000; Foellmi and Zweimüller 2006; García-Peñalosa and Wen 2008; Chu and Cozzi 2018).
2 Theoretical setup

In this section we extend the canonical quality-ladder, innovation-led growth model developed by Grossman and Helpman (1991) to allow for physical capital accumulation and two types of agents, namely workers and capitalists (see the online Appendix Section A for a complete derivation of the model). The model is used to organize the empirical analysis in the next sections, particularly, the functional relationship between the \( W/Y \) ratio, growth and investment in tangible and intangible assets.

2.1 The model

There is a homogeneous final good, \( \Upsilon \), which is taken as the numeraire, that can either be consumed or accumulated as physical capital. The good is produced by fully competitive firms using labour, capital and a continuum of intermediate inputs indexed by \( \omega \in [0, 1] \). The production function in the final good sector is

\[
\Upsilon = A_\Upsilon L_\Upsilon^{1-\gamma-\nu} K^\gamma D^\nu \quad \text{with} \quad 0 < \gamma, \nu \quad \text{and} \quad \gamma + \nu < 1,
\]

where \( A_\Upsilon \) is a constant that reflects the choice of units, \( L_\Upsilon \) is the amount of labour used in the production of final output, and \( K \) is physical capital. \( D \) is an index of intermediate inputs, which is defined as:

\[
\log D_t \equiv \int_0^1 \log \left[ \sum_j q_j(\omega)d_{jt}(\omega) \right] d\omega,
\]

where \( d_{jt}(\omega) \) denotes the input of the intermediate good \( \omega \) of quality \( j \) at time \( t \), and \( q_j(\omega) \) is its quality level. The quality of intermediate inputs increases over time due to technical progress. Normalizing the quality of each input to one at time \( t = 0 \), the quality \( j \) of product \( \omega \) amounts to \( q_j(\omega) = \lambda^j \), where \( \lambda > 1 \) is the size of the quality improvement that an innovation brings. In the intermediate-good sector, labour is the only primary factor of production. Regardless of product quality and product variety, one unit of labour is required to produce one unit of intermediate good. Thus, each firm has a marginal cost of production that is constant and equal to the real wage rate, \( w \). Producers of intermediate goods compete in prices. As a result, in each industry there is a leader that sells the state-of-the-art quality product, earning temporary monopoly profits that last until it is replaced by a new innovator.

Population consists of two types of infinitely-lived agents. Their lifetime utility function is given by

\[
U^j \equiv \int_0^\infty e^{-\rho t} \log C_t^x dt,
\]

where \( x \in \{k, l\} \). \( C^k \) is the consumption of a representative capital owner, whereas \( C^l \) is the consumption of a representative worker. All agents have the same discount rate \( \rho \). The number of each type of agent is constant over time and normalized to one. Only the capital owner accumulates tangible and intangible capital and maximizes utility subject to the following asset-accumulation equation:

\[
\dot{v} + \dot{K} = rv + (R - \delta)K - C^k,
\]
where $v$ is the real value of shares in intermediate-good firms, $r$ is the real interest rate, and $R - \delta$ is the real rental price of a unit of physical capital net of capital depreciation. The representative worker is endowed with one unit of labour, which is inelastically supplied.

Quality leaders in each industry are challenged by followers that engage in innovation races to discover the state-of-the-art quality product. All firms have the same innovation technology and there is free entry into the innovation race. Any innovating firm that invests $\chi \tilde{\iota}$ units of labour invents the next higher quality product with instantaneous probability $\tilde{\iota}$. Free entry into innovation races requires that the cost of one unit of labour employed in the innovation sector, $w$, is equal to its benefits, represented by the marginal product of labour in innovation, $1/\chi$, times the prize for a successful innovation, $v$, which gives $v = w\chi$.

The market-clearing condition in the labour market requires that demand for labour devoted to innovation and intermediate and final goods production equals labour supply. Aggregate consumption is $C \equiv C^k + C^l$, and investment demand equals $\dot{K} + \delta K$, where $\dot{K}$ represents net investment in the accumulation of physical capital. Thus, the market-clearing condition for the final good requires $\dot{K} + \delta K + C = \Upsilon$. Finally, two no-arbitrage conditions are used to complete the model. The first condition applies to the returns on equity claims and equates the sum of the firm’s dividend yield and the expected rate of capital gain to the interest rate. The second condition requires that the net rate of return to a unit of physical capital equals the interest rate.

The forces driving innovation fully determine the pace of economic expansion in the long run. On the balanced growth path, the wage rate, physical capital, consumption and final output all grow at a rate that is proportional to the aggregate arrival rate of innovation, $\iota$, namely

$$g_T = \left( \frac{\nu \log \lambda}{1 - \gamma} \right) \iota.$$  \hspace{1cm} (1)

2.2 The distribution of income in the economy and the wealth-income ratio

Capital owners hold their wealth in the form of equities in the intermediate-good firms and claims on physical capital. As the economy has a unit continuum of intermediate-good industries and there is a leader in each industry, the aggregate value of shares in the intermediate-good firms is equal to $v$. Thus, the aggregate stock of assets, $W$, amounts to $K + v$. Gross domestic product (GDP), $Y$, represents the economy’s total value added, which equals the sum of total output of the final good and the value created in the innovation sector, namely $Y + v$. As shown in the online Appendix Section A, GDP is equal to the sum of compensation to employees and gross operating surplus. Equivalently, GDP can be expressed as the sum of consumption and investment (savings), viz:

---

6The dynamic path of the economy is described by three differential equations, namely the two no-arbitrage conditions and the market-clearing condition for the final good. The analysis of this system of equations is complex, which justifies our decision to concentrate on the steady-state properties of the model, as we do in the empirical section. See the online Appendix Section A for further details.
\[ Y = Y + v = C + w\chi + \dot{K} + \delta K. \]

Here, we denote net national income as total income net of depreciation, \( Y - \delta K \). Furthermore, \( s_I \equiv \frac{w\chi \iota}{Y - \delta K} \) is the proportion of national income devoted to investment in intangible fixed assets, and \( s_T \equiv \frac{\dot{K}}{Y - \delta K} \) is the proportion of national income devoted to investment in tangible fixed assets. Thus, the \( W-Y \) ratio, \( \beta \), is the ratio between the aggregate stock of assets and net national income

\[ \beta \equiv \frac{W}{Y - \delta K} = \frac{K + v}{Y - \delta K}. \]

The steady-state value of \( \beta \) can be determined by decomposing \( \beta \) as the sum of \( \frac{K}{Y - \delta K} \) and \( \frac{v}{Y - \delta K} \). As capital and final output grow at the same rate in the steady state, \( \frac{K}{Y - \delta K} \) can be written as

\[ \frac{K}{Y - \delta K} = \frac{\dot{K}}{Y - \delta K} = \frac{s_T}{g_K} = \frac{s_T}{g_T}. \]

Using the free-entry condition \( v = w\chi \) and expressing the rate of innovation as \( \iota = g_T (1 - \gamma) / (\nu \log \lambda) \), we can write \( \frac{v}{Y - \delta K} \) as

\[ \frac{v}{Y - \delta K} = \frac{w\chi}{Y - \delta K} = \frac{1}{\iota} \cdot \frac{w\chi \iota}{Y - \delta K} = \frac{s_I}{\iota} = \frac{s_I}{g_T} \cdot \frac{\nu \log \lambda}{1 - \gamma}. \]

Therefore, the wealth-income ratio, \( \beta \), can be expressed as

\[ \beta = \frac{s_T}{g_T} \cdot \frac{s_I}{g_T} \cdot \Gamma, \quad (2) \]

where \( \Gamma \equiv \frac{\nu \log \lambda}{1 - \gamma} \). Eq. (2) adds a Schumpeterian (intangibles-driven) component of wealth accumulation to the steady state condition of the Solow growth model. This equation shows that \( s_T \) and \( s_I \) are both positively related to the \( W-Y \) ratio. However, while there is a one-to-one relationship between \( \beta \) and the \( s_T/g_T \) ratio, the relation between \( \beta \) and the \( s_I/g_T \) ratio depends on \( \Gamma \). Here, \( \Gamma \) is positively related to the innovation size, \( \lambda \), which essentially measures the extent to which higher quality inputs improve lower quality inputs. Thus, an invention that has a high commercial value because it significantly improves the quality of the product is capitalized at a higher rate than a lower value innovation.\footnote{In our growth setting, the relationship between innovation and income inequality is ambiguous (see Chu and Cozzi, 2018 for a similar result). From a theoretical standpoint, there is no perfect consensus regarding the sign of this relationship. Whereas Aghion et al. (2019) indicate that top income inequality is positively correlated with innovation, Jones and Kim (2018) find that the relationship goes the other way.}
3 Empirics: Model specification, data and causality

3.1 Model specification and estimation strategy

Our key equation, Eq. (2), predicts that the W-Y ratio, $\beta$, is positively related to the share of domestic income spent on intangible and tangible capital assets, $s_I$ and $s_T$, and is negatively related to the rate of economic growth, $g$, where all variables are expressed in net terms. We test this prediction by estimating the following log-linear model for 21 OECD countries over the period 1860-2015:

$$\ln \beta'_{it} = \eta_0 + \eta_1 \ln s'_{I,it} + \eta_2 \ln s'_{T,it} + \eta_3 \ln g'_{it} + \epsilon_{it}, \quad (3)$$

where $\eta_1 > 0$, $\eta_2 > 0$ and $\eta_3 < 0$, and $i$ refers to country $i$, and $t$ to time period. In our empirical specification, the variables are expressed in gross terms (denoted by a prime). Following Krusell and Smith (2015), income growth is defined as the growth in net domestic product plus the depreciation rate of total capital (tangible and intangible assets) denoted by $\delta$: $g' = g + \delta$. $\delta$ changes over time with the changing composition of the capital stock. Following the predictions of the model and the conventions of the literature (see, e.g., Madsen, 2010), the numerator of $s_T'$ is measured as non-residential (productive) investment.

We estimate the empirical model by means of the cross-sectionally augmented distributed lag (CS-DL) approach (Chudik et al., 2016). This allows us to identify the long-run impact of the regressors ($\eta$’s), by augmenting Eq. (3) with contemporary and lagged values of first-differenced regressors. Considering the long-run relationship between two variables

$$y_{it} = \eta \ x_{it} + \epsilon_{it},$$

the CS-DL specification takes the following shape:

$$y_{it} = \eta \ x_{it} + \sum_{\ell=0}^{p} \vartheta_{\ell} \Delta x_{it-\ell} + \omega_{it} \ y_{it} + \sum_{\ell=0}^{p} \omega_{i\ell} \ x_{it-\ell} + \epsilon_{it}, \quad (4)$$

where time-varying cross-sectional averages of the variables, denoted by bars, are included to control for cross-sectional error dependence induced by unobserved common shocks. These terms are known as common correlated effects, CCE (Chudik and Pesaran, 2015). While time dummies (TD) control for the effects of global shocks that affect all countries equally, the impact of CCE is country specific, and hence, these terms account for the spatially heterogeneous effects of un-observable factors, such as regional knowledge spillovers, military conflicts between neighbouring countries, etc. If these effects are not allowed for, the resulting cross-sectional error dependence will lead to inconsistent estimates in the event that un-observable common factors are correlated with the explanatory variables. An important

**Expressing Eq. (2) in logs enables us to use linear estimators, and mutes the influence of extreme observations (applying, particularly, to income growth). Note that all series have been re-scaled by adding the absolute value of the minimum of each variable to ensure that all observations are positive before logs are taken.**
property of Eq. (4) is that the first-difference terms capture the dynamic adjustment towards the steady state. This ensures that the coefficients of $s'_T$, $s'_I$, and $g'$ identify the steady-state relationship predicted by Eq. (2) and, therefore, are not biased by the influence of transitional dynamics.\footnote{As in Chudik et al. (2017), the number of lags, $p$, is set to the closest integer to $T^{1/3}$, where $T$ is the time dimension of the panel sample (for instance $p = 5$ in the regression using data from 1860). Following Eberhardt et al. (2013), we remove the effect of weak cross-sectional dependence (time dummies) by working with cross-sectionally demeaned variables.}

The main advantage of the CS-DL estimator is that it yields consistent long-run parameter estimates in a dynamic setting without being subject to the Nickell bias as may be the case for the auto-regressive distributed lag (ARDL) model. The CS-DL estimator has been shown to have good small sample properties in an array of conditions, such as, for instance, with stationary/non-stationary variables (and unknown factors), dynamic mis-specification, residual serial correlation, breaks in the error structure, etc. (Chudik et al., 2016; Chudik et al., 2017, p. 146).

### 3.2 Data

The analysis is based on a dataset covering the period 1860-2015 for the following 21 OECD countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the UK and the US. Tangible capital investment includes non-residential structures, machinery, equipment and weapon systems. Intangible capital investment includes expenditure on R&D, computer software, databases, and artistic originals. For the period 1960-2015, the data primarily come from OECD’s National Accounts.

We comply with the 2008 revision of the System of National Accounts (SNA) and include intangible assets or intellectual property products (henceforth IPP) in our estimates of investment and GDP, both in nominal and real terms. Intangible expenditure was not included as a part of GDP or investment in the earlier versions of national accounts as it was treated as intermediate input expenditure (Corrado et al., 2005; Koh et al., 2018). Accordingly, historical R&D expenditure is added to the official series on GDP and investment before the earliest year at which IPP data become available. The stocks of tangible and intangible fixed capital are constructed from real investment series using the perpetual inventory method and applying a depreciation rate of 3% for non-residential structures, 17% for machinery and equipment, 15% for R&D and intangible assets, where alternative depreciation rates for intangibles are allowed for in the robustness section below.

The wealth-income ratio, $\beta'$, is measured as the sum of the real stocks of tangible and intangible capital divided by real GDP. In robustness checks (see Section 4.3), $\beta'$ is measured at replacement cost and also as the ratio between the value of the stock market capitalization and GDP, where stock market capitalization is measured as the capital stock multiplied by Tobin’s $q$, as estimated in Madsen and Davis (2006). We measure $s'_I$ as the share of GDP invested in gross intangible assets, $s'_T$ as the share of GDP invested in gross tangible assets, and $g'$ as the growth rate of net domestic income augmented with the depreciation rate of total capital, $g' = g + \delta$, where $\delta$ changes over time as the composition of aggregate
capital stock changes (non-residential structures, machinery & equipment, and IPP).

The construction of R&D expenditure at current and constant prices before circa 1960 has been a Herculean task because these data are only available for a few countries and over a limited time period. The data are collected from numerous sources including annual issues of statistical yearbooks, special governmental budget material, and parliamentary papers. We derive the ratio between R&D expenditure and GDP using four types of data in descending order: 1) R&D outlays for industry and government; 2) government spending on R&D through tertiary education institutions; 3) government expenditure on tertiary education; 4) gross enrolment rates in tertiary education. Thus, the philosophy underlying our data construction is that R&D performed by the tertiary education sector is proportional to its size (i.e., a constant fraction of university funding/costs), which is not far in outcome from the method used by statistical agencies. Following the Frascati manual, hours worked by academics are scaled up by on-costs, multiplied by wages and added to the costs of equipment and buildings (see the online Appendix Section B).

We provide various evidence to substantiate the quality of the R&D data. We show that, for the US, there is a very strong correlation between the intangible investment ratio constructed by the Bureau of Economic Analysis and our data in the overlapping period 1929-2015, and we find evidence of cointegration between these two series (see the online Appendix Section C.1). More importantly, we run our regression model over different time periods (before and after 1945) and find that the impact of the intangible investment ratio is always significantly positive (see Section 4.1). Our results are consistent with historical evidence suggesting that R&D was mainly conducted in the tertiary education sector before WWII and, particularly, before WWI. The evidence for the US points towards close collaboration between universities and the agricultural and mining sectors – sectors that were crucial for the US industrialization. In fact, most R&D used for business purposes in Germany and the US was contracted out to universities and very little R&D was carried out by business enterprises. In the first edition of Science in 1906, James McKeen Cattell, the editor, published 4,000 biographies of researchers in the US; however, no researchers from industry were included in the directory (Godin and Schauz, 2016, p. 279), suggesting that the share of private industry R&D in total R&D must have been miniscule at that time in the US and, presumably, in the other countries in our sample (see Arora et al., 2020). Furthermore, starting from a low base, the available data indicate a markedly growing share of R&D undertaken by the private industry after WWII (see the online Appendix Section B for discussion). Another challenge has been the construction of the R&D deflator. We compute it as the weighted average of salaries of R&D workers (60%), construction cost deflator (20%) and deflator for investment in machinery and equipment (20%), where the weights reflecting the composition of R&D outlays have stayed relatively constant over time (see, e.g., Igna and Venturini, 2019). Again, the data are derived from numerous national statistical sources. Since data on the remuneration of researchers, scientists and technical personnel are rarely available, their wages have been mostly approximated by the remunerations of highly educated professionals or the average income transformed to earnings of high-income earners using information on top-income distributions. Full details on data sources, data construction and related robustness checks are provided in the online
Appendix Sections B and C.

4 Regression results

We start our analysis by showing our baseline regressions with and without control variables and for different time periods in Section 4.1. Subsequently we test the sensitivity of the parameter estimates to an alternative estimation method that addresses simultaneity; the use of catastrophic events as external instruments; allowance for capital measured at market values and variation of capital depreciation rates; to restricting IPP capital to R&D only; estimation with five-year non-overlapping observations and heterogeneity in the regressors’ impact across time and space. A detailed discussion of the results in this section is relegated to the online Appendix Section C.

4.1 Baseline estimates

The results of regressing Eq. (3) are shown in Table 1. Column (1) includes only the variables that, according to our theoretical setting, should be related to the steady-state value of the $W/Y$ ratio. In line with the predictions of our theoretical framework, the coefficients of $s'_I$ and $s'_T$ are significantly positive, and the coefficient of $g'$ is significantly negative. A one percent increase in $s'_I$ ($s'_T$) results in a 0.15 (0.48) percent increase in the $W/Y$ ratio, whereas a one percent increase in the rate of income growth results in a 0.03 percent reduction in the $W/Y$ ratio. Note that the effect of the investment ratio on the $W/Y$ ratio is higher for $s'_T$ because its share in total GDP over the period 1860-2015 is well above that of $s'_I$.

In columns (2)-(4), we assess the stability of the coefficients of the explanatory variables across the estimation periods 1860-1945, 1945-2015 and 1970-2015, where 1945 signifies a structural shift in the growth regime (Greasley et al., 2013), while the post-1970 period has been characterized by increasing inequality (Piketty and Zucman, 2014), the transition to a regime of reduced productivity growth, the ICT revolution, and the markedly increased investment in IPP in the last few decades (Cette et al., 2016; Crafts, 2018; Koh et al., 2018). The stability of the coefficients on an annual basis is discussed below in Section 4.3. The coefficients of the investment ratios remain statistically significant across estimation periods and their magnitudes are consistent with the changing composition of aggregate capital through time. The coefficient of the tangible investment rate, $s'_T$, fluctuates around 0.5 throughout the entire period, while the coefficient of $s'_I$ has increased over time. Though lower than estimated over the full time horizon, it is noteworthy that the coefficient of the intangible investment ratio is significant in the pre-WWII period, suggesting the relatively good quality of our historical data on IPP. Finally, the absolute value of the coefficient of income growth, $g'$, has increased markedly over time, which is unsurprising since the rate of economic expansion was quite volatile in the past due to the two world wars, two depressions during the interwar period and one in the 1890s, yielding a high noise-to-signal ratio in this variable.

Piketty and Zucman (2014), Roine and Waldenström (2015) and Piketty (2015b) suggest that the $W/Y$ ratio was pulled off of its steady state during the approximate period 1920-1980 due to war destruction,
Table 1: CS-DL estimates of the W-Y ratio (1860-2015)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intangibles invest /GDP</td>
<td>0.153*** (0.024)</td>
<td>0.048* (0.027)</td>
<td>0.039* (0.023)</td>
<td>0.071*** (0.025)</td>
<td>0.151*** (0.023)</td>
<td>0.153*** (0.024)</td>
<td>0.153*** (0.025)</td>
<td>0.154*** (0.025)</td>
</tr>
<tr>
<td>Tangibles invest /GDP</td>
<td>0.477*** (0.033)</td>
<td>0.407*** (0.041)</td>
<td>0.565*** (0.048)</td>
<td>0.504*** (0.057)</td>
<td>0.426*** (0.037)</td>
<td>0.479*** (0.033)</td>
<td>0.480*** (0.035)</td>
<td>0.471*** (0.032)</td>
</tr>
<tr>
<td>Income growth</td>
<td>-0.031*** (0.006)</td>
<td>-0.016*** (0.004)</td>
<td>-0.629*** (0.091)</td>
<td>-0.958*** (0.083)</td>
<td>-0.031*** (0.007)</td>
<td>-0.031*** (0.006)</td>
<td>-0.031*** (0.006)</td>
<td>-0.031*** (0.006)</td>
</tr>
<tr>
<td>Financial development</td>
<td>0.102*** (0.021)</td>
<td>0.102*** (0.013)</td>
<td>0.102*** (0.017)</td>
<td>0.102*** (0.017)</td>
<td>0.102*** (0.017)</td>
<td>0.102*** (0.017)</td>
<td>0.102*** (0.017)</td>
<td>0.102*** (0.017)</td>
</tr>
<tr>
<td>Inflation rate</td>
<td>0.003 (0.017)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
</tr>
<tr>
<td>Overall tax rate</td>
<td>-0.030** (0.013)</td>
<td>-0.030** (0.013)</td>
<td>-0.030** (0.013)</td>
<td>-0.030** (0.013)</td>
<td>-0.030** (0.013)</td>
<td>-0.030** (0.013)</td>
<td>-0.030** (0.013)</td>
<td>-0.030** (0.013)</td>
</tr>
<tr>
<td>Trade openness</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
<td>0.081* (0.049)</td>
</tr>
<tr>
<td>Observations</td>
<td>3,038</td>
<td>1,610</td>
<td>1,386</td>
<td>882</td>
<td>3,038</td>
<td>3,038</td>
<td>3,038</td>
<td>3,038</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.501</td>
<td>0.664</td>
<td>0.362</td>
<td>0.287</td>
<td>0.470</td>
<td>0.497</td>
<td>0.501</td>
<td>0.493</td>
</tr>
</tbody>
</table>

Notes: CS-DL estimates. The dependent variable is the gross capital-income ratio ($\beta'$). Intangible assets are defined as total IPP. Standard errors are in parentheses and the variables are measured in logs. The reported coefficients are the long-run parameters. All regressions include country-specific fixed effects and CCE. Financial development is measured as the ratio of bank credit to GDP (col. (5)); the inflation rate is measured as the rate of change in the Consumer Price Index (col. (6)); overall tax rate is measured as the ratio of overall tax revenues to GDP (col. (7)); trade openness is measured as the ratio of imports to GDP (col. (8)). ***, **, * significant at 1, 5 and 10%.

Inflation that eroded the real value of bonds, and shocks to expected post-tax returns to capital, such as tax hikes and credit constraints. If these variables are correlated with the focus regressors, but omitted from the regressions, the parameter estimates of investment rates and income growth might be biased. To deal with potentially relevant omitted variables, we include financial development, consumer price inflation, the overall tax rate and trade openness as controls.\(^{10}\)

The regression in column (5) includes the share of bank credit in total GDP as a proxy for financial development. This is a potentially important control variable because financial development simultaneously impacts investment in knowledge production (mainly R&D) and investment in tangible assets \(^{11}\). The coefficient of financial development is significantly positive, suggesting that financial development may lead to a better evaluation of capital assets through more efficient intermediation. Our principal results are unaffected by the inclusion of financial development in the regression.

The consumer price inflation rate is included in the regression in column (6). As stressed by Piketty \(^{12}\), inflation erodes the real value of wealth and, at the same time, reduces investment because it heightens the expected profitability of future investment projects and, consequently, increases the option value of postponing investment. Furthermore, inflation reduces the capitalized real value of depreciation

---

\(^{10}\)Financial development is measured as the share of credit to the non-financial private sector in GDP; the inflation rate is measured as the rate of change in consumer prices; the tax rate is measured as the share of the tax revenues in GDP; trade openness is measured as the import-GDP ratio. See the online Appendix Section B for data sources.
of fixed capital for tax purposes. In our regression, the coefficient of inflation is significantly negatively related to the $W-Y$ ratio as expected; however, the baseline results remain intact.

In the regression in column (7) we control for the overall tax rate, computed as the share of overall tax revenues in GDP. As shown by Gemmell et al. (2011, 2014), taxation impacts investment, factor productivity and economic growth; factors that are related to the $W-Y$ ratio. The coefficient of the overall tax rate is insignificant in the regression in column (7), while the coefficients of $s_I'$ and $g'$ remain close to those of the baseline regression.

Trade openness is included in the regression in column (8). Theoretically, the sign of trade openness is ambiguous since trade openness may promote investment as well as income growth (Prankel and Romer, 1999); thus simultaneously affecting the numerator and the denominator of the $W-Y$ ratio. Trade openness may also favour the world transmission of technological ideas (Madsen, 2007). Furthermore, trade openness may affect both factor income and the distribution of personal income through offshoring practices and imports from low-income countries that are exporting unskilled labour intensive products. However, evidence on this linkage is still unclear (Elsby et al., 2013; Autor et al., 2017; Cerdeiro and Komaromi, 2017). In column (8), while the coefficients of the key variables remain consistent with those of the baseline regression, the coefficient of trade openness is significantly positive at a 10% level, suggesting that increasing trade openness contributes somewhat to wealth inequality.

In sum, our baseline results remain robust to the factors that may have influenced the evolution of the $W-Y$ ratio in our sample period but that are not accounted for in our theoretical framework.

### 4.2 Simultaneity issues

Although the CS-DL estimator provides consistent estimates under some regularity conditions, it is not robust to feedback effects from the dependent variable to the regressors; thus potentially resulting in biased coefficient estimates. To tackle this type of endogeneity, we carry out two alternative types of regressions, namely ARDL and IV regressions (see the online Appendix Section C.2 for further details). The ARDL estimator provides consistent parameter estimates in the presence of feedback effects from the dependent variable when the lag structure of the variables is correctly specified, regardless of their order of integration (Chudik et al., 2016; Chudik et al., 2017, p. 143). The IV estimator has the advantage of exploiting exogenous variation in a set of factors that are external to our conceptual framework; however, this identification strategy is more conventional for the static analysis.

In the IV regressions, we exploit exogenous variation in the occurrence of rare natural disasters to predict the effect of our explanatory variables. The economic literature on natural disasters provides two main insights that are useful for our identification strategy (Cavallo and Noy, 2009). First, in the aftermath of a catastrophic event, the level of economic activity is reduced as a consequence of the catastrophe in the affected area. The catastrophic event creates incentives to replace destroyed capital stock and to improve technologies that favour long-run growth prospects of the affected areas (Skidmore and Toya, 2002; Mia and Popp, 2014). Consequently, the short- and long-run effects of catastrophic events may differ as demand conditions and incentives to invest change over time. Second,
the economic effects of natural disasters propagate spatially (Xiao and Nilawar, 2013). Due to the
destruction and production interruptions, local demand re-orientates towards goods and technologies
produced in other countries and this may speed up economic growth in the unaffected areas (Gassebner
et al. 2010, Cuaresma et al. 2008). For instance, Ager et al. (2019) document that the 1906 San Francisco
Earthquake permanently affected the distribution of economic activity within the US promoting workers’
migration across states and raising manufacturing production in the cities located in unaffected areas.

We use geographic distance proximity-weighted cumulated occurrences, in a window of up to six
years, of earthquakes and extreme temperatures as instruments for $s'_I$, volcanic eruptions and wildfires
for $s'_T$, and droughts and floods for $g'$. These events are exogenous and are likely to be correlated with
domestic economic variables through spatial propagation of the effects of disasters. Geographic proximity
is used instead of trade flows since trade may be affected by the spatial propagations of the disasters (see,
e.g., Gassebner et al. 2010 and Pelli and Tschopp, 2017 for evidence on trade responsiveness to natural
catastrophic events). The cumulated frequency of natural disasters is used instead of proxies for the
disaster’s magnitude, such as the scale of damage or the number of injured/affected people, as these would
be influenced by the level of economic development achieved by the affected areas before the catastrophe.
We use data on natural disasters from EM-DAT, the International Disaster Dataset, managed by the
Centre for Research on the Epidemiology of Disasters (CRED), at the Université Catholique de Louvain
(Belgium). This includes information on the number of mass disaster events, the estimated damage and
the approximate number of people affected. CRED collects data from different sources and provides
consistent series starting from 1900. Following Bloom et al. (2013), the effect of each endogenous variable
is estimated by means of a set of first-stage (static) regressions; in the second stage, these predicted values
are used as explanatory variables in our CS-DL regressions.

The first-stage results of such regressions are shown in Table C1 of the online Appendix Section C.2.
The coefficients of natural disasters have their expected signs and the $F$-tests of exclusion restrictions are
all in the region 10.5-17.2, suggesting that the relevance criterion is satisfied. The second-stage results,
with and without control variables, are shown in Table 2. The coefficients of our focus regressors are
statistically significant in all cases when they are instrumented, and regardless of whether control variables
are included in the regressions. The coefficients of the focus variables, tangible and intangible investment
ratio and income growth, are close to those of the baseline CS-DL estimates in Table 1, suggesting that
the baseline results are not biased or inconsistent because of feedback effects from the $W$-$Y$ ratio and
omission of unobserved important variables.

The ARDL regression results are reported in Tables C1 (without controls) and C2 (with controls). The
coefficients of the focus variables are all significant at the 1 percent level and the coefficients are con-

11 Both stages of IV regressions are estimated using country-specific fixed effects and common correlated effects.
The second-stage regressions are based on bootstrapped standard errors with 200 replications.
12 We estimate the following cross-sectionally augmented ARDL specification:

\[ y_{it} = \sum_{\ell=0}^{5} \vartheta_{x,\ell} x_{it-\ell} + \sum_{\ell=1}^{5} \vartheta_{y,\ell} y_{it-\ell} + \sum_{\ell=0}^{5} \vartheta_{z,\ell} z_{it-\ell} + \sum_{\ell=0}^{5} \vartheta_{\eta,\ell} \eta_{it-\ell} + \epsilon_{it}. \]

In Tables C1-C3 in the online Appendix Section C.2 we report the long-run coefficients, obtained as

\[ \eta = \sum_{\ell=0}^{5} \vartheta_{x,\ell} / (1 - \sum_{\ell=1}^{5} \vartheta_{y,\ell}), \]

and assess the significance of these parameters using the delta method.
sistent with those of the CS-DL regression assuming exogenous regressors (Table 1), or that instrumenting regressors with natural disasters (Table 2). Furthermore, we check whether the ARDL and CS-DL results are robust to the lag structure, noting that a sufficiently long number of lags is a requisite for getting consistent ARDL estimates, while a too large number of lags may reduce the estimation efficiency. Conversely, the CS-DL estimates are less sensitive to lag specification than ARDL estimates (Chudik et al., 2017, p. 144). The results, which are displayed in Table C3, show that long-run coefficients are unaffected by the lag structure of the variables. All the coefficients of the focus variables are significant at the 1 percent level and are close of those of the baseline regressions, suggesting that the baseline regressions are unlikely to be significantly influenced by feedback effects from the dependent variable.

4.3 Sensitivity analysis

This sub-section summarizes the results of the robustness checks that are reported and discussed in depth in the online Appendix Sections C.3-C.5.

Measurement issues. Since our analysis goes back to almost the mid-19th century, we check for the sensitivity of the estimates to changes in the way the variables are measured.

In Table C4, we first measure $\beta'$ at replacement costs by multiplying the real capital stock and GDP by their own deflators. Whilst the coefficients of $s'_I$ and $g'$ remain approximately unchanged relative to the baseline results, the coefficient of the tangible investment ratio, $s'_T$, has increased; thus highlighting the importance of measuring $\beta'$ in real terms. Next, we restrict our measure of intangibles to R&D; the coefficients remain close to baseline estimates. Finally, we measure the $W/Y$ ratio as the ratio between the stock market capitalization of non-residential capital and income. Thus, wealth is measured in market values as opposed to acquisition costs, coming close to the market price measures of wealth used in Piketty and Zucman (2014), Piketty (2015b), and Madsen (2019). However, note that the measure of wealth based on stock market values is not consistent with our theoretical set-up in which Tobin’s $q$ is equal to one in steady state. In this estimation, the coefficient of $g'$ is no longer significant, presumably because the employed measure of wealth is more sensitive to bubbles and financial speculation, and is based less on economic fundamentals.

In Table C5 we assess the sensitivity of the estimates to variations in the depreciation rates of intangible knowledge stock. The evidence suggests that the rate of depreciation of technological knowledge has increased over time due to the changing composition of R&D and the pace of scientific advancement; particularly the increasing share of software in intangible investment (Li and Hall, 2020). To cater for this, we estimate the baseline model using various depreciation rates for intangibles: (i) constant 10% and 40% depreciation rates throughout the entire estimation period; and (ii) time-varying depreciation rates starting from 10% in 1860 and growing linearly or exponentially to 40% in 2015. Regardless of which of these depreciation rates are used, the coefficient of $s'_I$ and the other explanatory variables remain close to those of the baseline regression, suggesting that our baseline results are not sensitive to the choice of depreciation rates.
Table 2: CS-DL IV regressions including controls (1900-2015)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section A - instrumented</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intangible invest/GDP</td>
<td>s_I'</td>
<td>0.203***</td>
<td>0.202***</td>
<td>0.186*</td>
<td>0.205**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.092)</td>
<td>(0.072)</td>
<td>(0.100)</td>
<td>(0.091)</td>
</tr>
<tr>
<td>Tangible invest/GDP</td>
<td>s_T'</td>
<td>0.374***</td>
<td>0.368***</td>
<td>0.372***</td>
<td>0.375***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.066)</td>
<td>(0.073)</td>
<td>(0.071)</td>
<td>(0.089)</td>
</tr>
<tr>
<td>Income growth</td>
<td>g'</td>
<td>-0.020*</td>
<td>-0.020**</td>
<td>-0.021</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.010)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Financial development</td>
<td></td>
<td>-0.010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.033)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflation rate</td>
<td></td>
<td></td>
<td>-0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.033)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall tax rate</td>
<td></td>
<td></td>
<td>-0.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.078)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade openness</td>
<td></td>
<td></td>
<td>0.037</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.080)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.618</td>
<td>0.623</td>
<td>0.616</td>
<td>0.608</td>
<td>0.622</td>
</tr>
</tbody>
</table>

| **Section B - instrumented** |             |             |             |             |             |
| Intangible invest/GDP | s_I'        | 0.181***    | 0.151***    | 0.145***    | 0.180***    | 0.188***    |
|                      |             | (0.035)     | (0.040)     | (0.051)     | (0.030)     | (0.030)     |
| Tangible invest/GDP  | s_T'        | 0.413***    | 0.333**     | 0.367***    | 0.362***    | 0.416***    |
|                      |             | (0.127)     | (0.143)     | (0.113)     | (0.113)     | (0.095)     |
| Income growth        | g'          | -0.023      | -0.026      | -0.022      | -0.021      | -0.025      |
|                      |             | (0.023)     | (0.017)     | (0.021)     | (0.028)     | (0.023)     |
| Financial development|             | 0.106**     |             |             |             |             |
|                      |             | (0.053)     |             |             |             |             |
| Inflation rate       |             |             | -0.009      |             |             |             |
|                      |             |             | (0.046)     |             |             |             |
| Overall tax rate     |             |             | 0.004       |             |             |             |
|                      |             |             | (0.058)     |             |             |             |
| Trade openness       |             |             | 0.007       |             |             |             |
|                      |             |             | (0.094)     |             |             |             |
| R-squared            | 0.700       | 0.690       | 0.724       | 0.705       | 0.694       |

| **Section C - instrumented** |             |             |             |             |             |
| Intangible invest/GDP | s_I'        | 0.127***    | 0.121***    | 0.112***    | 0.132***    | 0.136***    |
|                      |             | (0.025)     | (0.019)     | (0.030)     | (0.030)     | (0.026)     |
| Tangible invest/GDP  | s_T'        | 0.411***    | 0.423***    | 0.398***    | 0.445***    | 0.399***    |
|                      |             | (0.096)     | (0.102)     | (0.086)     | (0.087)     | (0.106)     |
| Income growth        | g'          | -0.068*     | -0.073**    | -0.071*     | -0.074**    | -0.078**    |
|                      |             | (0.037)     | (0.029)     | (0.037)     | (0.033)     | (0.036)     |
| Financial development|             | 0.001       |             |             |             |             |
|                      |             | (0.073)     |             |             |             |             |
| Inflation rate       |             |             | -0.029      |             |             |             |
|                      |             |             | (0.031)     |             |             |             |
| Overall tax rate     |             |             | -0.059      |             |             |             |
|                      |             |             | (0.056)     |             |             |             |
| Trade openness       |             |             | 0.010       |             |             |             |
|                      |             |             | (0.123)     |             |             |             |
| R-squared            | 0.593       | 0.583       | 0.613       | 0.588       | 0.590       |
| Intangible assets    | IPP         | IPP         | IPP         | IPP         | IPP         |
| Obs                  | 2,205       | 2,205       | 2,205       | 2,205       | 2,205       |

**Notes:** CS-DL estimates with IV. The dependent variable is the gross capital-income ratio ($\beta'$). Intangible assets are defined as total IPP. Standard errors are in parentheses (bootstrapped with 200 replications). Variables are measured in logs. The reported coefficients are the long-run parameters. All regressions include country-specific fixed effects and CCE. Financial development is measured as the ratio of bank credit to GDP; the inflation rate is measured as the rate of change in the Consumer Price Index; overall tax rate is measured as the ratio of overall tax revenues to GDP; trade openness is measured as the ratio of imports to GDP. ***, **, * significant at 1, 5 and 10%.
To rule out the possibility of any estimation bias induced by short-run co-movements between the variables at business-cycle frequencies, we estimate the baseline regression using observations in non-overlapping 5-year intervals. Regardless of whether the models are specified dynamically or statically, the results are close to the CS-DL estimates at annual frequencies, especially the parameter estimates of the investment ratios, \( s'_I \) and \( s'_T \) (see Table C6).

**Parameter heterogeneity.** As a final check on the empirical model, we allow the coefficients of the explanatory variables of our baseline regression to vary across countries and over time, using the estimation procedures devised by Chudik and Pesaran (2015) and Neal (2018). The idea is to explore whether our baseline results are influenced by parameter heterogeneity along both dimensions of the panel sample (time and space). The results, which are reported in Table C7, show that the coefficients of \( s'_I \) and \( s'_T \) are unaffected by the allowance of parameter heterogeneity in the time or space dimensions, or both. The coefficient of \( g' \) turns out to be statistically and economically highly significant in the regression that allows for country-specific heterogeneity. However, the coefficient of income growth becomes less stable when allowing for parameter heterogeneity, both in the cross-sectional and in the time dimensions, probably reflecting the alternation of periods of economic expansion (\( g' \) positive) and recession (\( g' \) negative).

5 Quantifying the net effect of intangibles on wealth inequality

In this section, we quantify the portion of wealth inequality that can be explained by our model. We restrict our investigation to a few drivers of the wealth-income ratio following the predictions of our theoretical framework. This exercise is tentative as we may neglect some important determinants identified by the literature, which we have not controlled for in the regressions in Tables 1 and 2.

---

13 Our analysis is restricted in three ways. First, we may have omitted important variables. Of potentially important determinants of the \( W-Y \) ratio not controlled for are, according to Acemoglu and Robinson (2015), institutions and political factors. The increasing concentration of capital in the US after the Civil War, for example, sparked a popular movement against the concentration and led to the introduction of regulatory laws (Acemoglu and Robinson, 2015). Also, sectoral changes associated with economic development and skill-biased technological progress may be behind the change in the \( W-Y \) ratio by promoting the allocation of resources towards more innovation-intensive sectors, away from physical capital-intensive activities. Second, we do not distinguish between different skill levels. Our setup considers only one representative worker and, therefore, there is no occupational choice between employment in production or in R&D. Our analysis, therefore, abstracts from the fact that workers may occupy positions of different skill levels at different wages. An extensive literature has analysed how agents’ occupational choices affect the long-run wealth distribution, wealth polarization and economic growth (see, among others, Banerjee and Newman, 1993; Galor and Zeira, 1993; Chakraborty and Citanna, 2005). Third, the model is restricted by the assumption that agents have equal investment opportunities. Consequently, they realize equal expected returns without facing idiosyncratic investment risk. The degree to which individual households are exposed to idiosyncratic risk may, however, positively affect the rate at which wealth concentrates at the top (Piketty, 2014). Omission of these factors may bias the estimates if they simultaneously influence the \( W-Y \) ratio and, at the same time, are correlated.
In the neoclassical growth framework, which is adopted by Piketty and Zucman (2014), Piketty (2015b) and Madsen (2019), an increase in the saving/investment ratio is unambiguously associated with a higher \( W/Y \) ratio, as the steady-state income growth rate is exogenous and, hence, unrelated to the investment ratio. In our Schumpeterian setting, the link between innovative investment and wealth inequality is not clear cut as R&D and other intangibles have two counterbalancing effects on the \( W/Y \) ratio. First, an increasing share of GDP invested in knowledge-generating activities is associated with a higher ratio between corporate wealth and national income, as the incentive to innovate is driven by the market value of the firms (wealth channel, WC). A higher innovation rate, resulting from a greater research engagement, increases the rents associated with intangible investments, thereby promoting a more uneven distribution of the resources in the economy. This is the direct effect of \( s_I \) on wealth inequality. Second, successful R&D leads to a higher rate of innovation, which spurs the rate of income growth and this ultimately reduces the \( W/Y \) ratio (growth channel, GC). This is the indirect effect of \( s_I \) on wealth inequality. The two channels can be formalized as follows

\[
\begin{align*}
  s_I' \to s_I' \Gamma + \frac{\bar{s}_T}{g'} \to \frac{W}{Y}, & \quad \text{Wealth channel (WC)} \\
(s_I' \to \iota \to) \quad g' \to s_I' \Gamma + \frac{\bar{s}_T}{g'} \to \frac{W}{Y}, & \quad \text{Growth channel (GC)}
\end{align*}
\]

where a bar over a variable means that this is kept constant. Consistent with the model, the rate of innovation, \( \iota \), is defined as the patenting rate.

In the regression analysis thus far, we have only focussed on the direct effects on innovations on the wealth inequality ratio by estimating the reduced-form equation for the \( W/Y \) ratio. However, to identify both mechanisms at work and to assess the net linkage between intangible assets and the \( W/Y \) ratio, we estimate the following equation system

\[
\begin{align*}
  \ln \iota_{it} &= \xi_0 + \xi_1 \ln s_{I, it} + \epsilon_{1, it}, \\
  \ln g_{it} &= \zeta_0 + \zeta_1 \ln \iota_{it} + \epsilon_{2, it}, \\
  \ln \beta_{it} &= \pi_0 + \pi_1 \ln s_{I, it} + \pi_2 \ln s_{T, it} + \pi_3 \ln g_{it} + \epsilon_{3, it}.
\end{align*}
\]

Each equation is modelled as a CS-DL specification, hence providing long-run effects of the variables. Following Crepon et al. (1998), we estimate the system by 3SLS and derive the total (net) effects of changes in \( s_I' \) on the \( W/Y \) ratio as follows:

\[
\frac{\partial \ln \beta'}{\partial \ln s_I'} = \frac{\partial \ln \beta'}{\partial \ln s_I'} \bigg|_{s_I'=s_I'} + \frac{\partial \ln \beta'}{\partial \ln s_I'} \bigg|_{g'=g'} = \frac{\partial \ln \beta'}{\partial \ln g'} \cdot \frac{\partial \ln g'}{\partial \ln \iota} \cdot \frac{\partial \ln \iota}{\partial \ln s_I'} + \pi_1 = \xi_1 \zeta_3 + \pi_1 = \xi_1 \zeta_3 + \pi_1, \quad (8)
\]

where GC is the growth effect and WC is the wealth effect of changes in \( s_I' \).

The system estimates are presented in Table 3 where the baseline estimates are displayed in column

\[14\] Each equation of the system uses one-year lagged regressors.
Table 3: System estimates and total (net) effect of the intangible investment ratio, $s'_I$

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bench-mark System</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dep. variable:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>$\iota$</td>
<td>$g'$</td>
<td>$\beta$</td>
<td>$\iota$</td>
<td>$g'$</td>
<td>$\beta$</td>
<td>$\iota$</td>
<td>$g'$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Intangibles invest/GDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s'_I$</td>
<td>0.209***</td>
<td>(0.044)</td>
<td>0.143***</td>
<td>(0.099)</td>
<td>0.058***</td>
<td>(0.018)</td>
<td>0.025*</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Rate of innovation</td>
<td>$\iota$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\iota$</td>
<td>0.265***</td>
<td>(0.090)</td>
<td>0.422***</td>
<td>(0.009)</td>
<td>0.587***</td>
<td>(0.018)</td>
<td>0.017</td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>Intangibles invest/GDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s'_I$</td>
<td>0.153***</td>
<td>(0.024)</td>
<td>0.143***</td>
<td>(0.010)</td>
<td>0.058***</td>
<td>(0.010)</td>
<td>0.017</td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>Tangibles invest/GDP</td>
<td>$s'_T$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s'_T$</td>
<td>0.477***</td>
<td>(0.033)</td>
<td>0.422***</td>
<td>(0.009)</td>
<td>0.587***</td>
<td>(0.018)</td>
<td>0.017</td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>Income growth</td>
<td>$g'$</td>
<td>-0.031***</td>
<td>(0.006)</td>
<td>-0.011</td>
<td>(0.011)</td>
<td>-0.588***</td>
<td>(0.061)</td>
<td>-0.717***</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Obs.</td>
<td>3,038</td>
<td>3,075</td>
<td>3,075</td>
<td>3,075</td>
<td>1,297</td>
<td>1,297</td>
<td>1,297</td>
<td>1,297</td>
<td>1,297</td>
</tr>
<tr>
<td>$R$-squared</td>
<td>0.502</td>
<td>0.668</td>
<td>0.131</td>
<td>0.925</td>
<td>0.848</td>
<td>0.955</td>
<td>0.869</td>
<td>0.955</td>
<td>0.971</td>
</tr>
<tr>
<td>Intangible assets</td>
<td>IPP</td>
<td>IPP</td>
<td>IPP</td>
<td>IPP</td>
<td>R&amp;D only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wealth channel</td>
<td>WC</td>
<td>0.143</td>
<td>0.058</td>
<td>0.045</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth channel</td>
<td>GC</td>
<td>-0.006</td>
<td>-0.014</td>
<td>-0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (net) effect of intangibles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.142</td>
<td>0.044</td>
<td>0.038</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: 3SLS estimates of Eqs. (5)-(7). Each equation is modelled as a CS-DL specification. Intangible assets are defined as IPP in cols. (1)-(7), and as R&D only in cols. (8)-(10). Standard errors are in parentheses. Variables are measured in logs. The reported coefficients are long-run parameters. All regressions include country-specific fixed effects and CCE. ***, **, * significant at 1, 5 and 10%.

(1). Estimates for the period 1860-2015 are shown in columns (2)-(4), and post-1945 estimates are presented in columns (5)-(10). Post-1945 regressions in which intangibles are measured by R&D only are presented in the last three columns of the table.

The parameter estimates are close to the CS-DL and IV estimates in Tables 1 and 2; thus giving further support to the baseline regression results. The estimated effects of the intangible investment ratio on the innovation rate, $\iota$, measured as the number of patent applications divided by population, are displayed in columns (2), (5) and (8). The coefficient of $s'_I$ is highly significantly positive in all regressions, suggesting that intangibles are the key drivers of innovation successes. The economic impact of intangibles is higher in the estimates covering the post-WWII period than the estimates covering earlier periods, probably because innovations in the pre-WWII period were less systematic and less targeted at industrial applications than those of the post-WWII period (Mokyr, 2018). Following the predictions of our theoretical framework (see Eq. 1) and other Schumpeterian growth models, the regressions in columns (3), (6) and (9) indicate a significantly positive relationship between the rate of innovation, $\iota$, and economic growth, $g'$. The finding that income growth is proportional to $s'_I$ in the long run and, therefore, that $s'_I$ has persistent growth effects is consistent with the evidence of Madsen (2010) and Minniti and Venturini (2017), and lends support to fully endogenous growth theory (Aghion and Howitt 1998; Peretto 1998, 1999; Howitt 1999; Peretto and Smulders 2002). Finally, the estimates of the direct drivers of the $W-Y$ ratio (Eq. 7) in columns (4), (7) and (10) are comparable to the baseline results of
Based on Eq. (8), the direct and indirect effects of the intangible investment ratio on the $W-Y$ ratio are shown in the lower panel of Table 3 (WC and GC). The total (net) effect of $s'_I$ is higher in the regression over the long time span (columns (2)-(4)) than in the post-1945 regressions (columns (5)-(7)), reflecting that the growth channel has become effective mainly in the second half of our time interval. Interestingly, after 1945 the total effect of $s'_I$ on wealth inequality is approximately similar regardless of whether intangibles are measured as IPP (0.044) or R&D (0.038).

Based on the estimates over the period 1945-2015, a 100% increase in $s'_I$ is associated with a 4% increase in the $W-Y$ ratio. While this overall effect may look small, it is noteworthy that the share of intangible investment in total GDP, as measured by IPP (R&D), has increased by almost 1200% (800%) since 1945 and has converged to the share of tangible investment, $s'_T$ (see Corrado et al., 2017). The 12-fold (8-fold) increase in $s'_I$ has resulted in a 53% (30%), increase in the $W-Y$ ratio, suggesting that the expansion of intangibles has been highly influential for the widening of wealth inequality in the post-WWII period, particularly from 1970 onwards (see Figure A.1 of the online Appendix Section C.1).

This result is consistent with the finding of Koh et al. (2018) that the decrease in labour’s income share in the US after WWII has been predominantly driven by the expansion of IPP capital. Tangible investment, by contrast, has had quite different effects on the inequality path. Based on a coefficient of $s'_T$ of 0.59, the two-fold increase in the tangible investment ratio observed from 1945 to 1963 would have raised the $W-Y$ ratio by almost 117%. Conversely, the 44% (33%) decline in $s'_T$ over the period 1964-2015 (1980-2015) contributed to a decrease in the $W-Y$ ratio by 26% (19%), suggesting that tangible capital has not been a source of the increasing $W-Y$ ratio during the last decades.

6 Conclusions

The increasing income and wealth inequality over the last decades in the OECD countries has been attributed to various factors, including a decline in relative prices of investment goods (Karabarbounis and Neiman, 2014), a deceleration in human capital accumulation (Grossman et al., 2017), an increase in the wealth-income ratio induced by reduced productivity growth (Piketty and Zucman, 2014), and an automation of tasks previously performed by labour (Acemoglu and Restrepo, 2018).

Along these lines, we contribute to this literature by showing that intangibles have been influential for wealth inequality since 1860 and they have been a key driver of inequality since the 1970s. In a Schumpeterian growth model, the $W-Y$ ratio is shown to be governed by both the tangible and intangible investment ratios and income growth in steady state. An implication of this set-up is that intangibles increase the $W-Y$ ratio by raising the capitalized value of the firms and, at the same time, they reduce the $W-Y$ ratio through innovation-driven growth. The model goes a step further than the neoclassical growth model by making a clear distinction between tangible and intangible investments and by endogenizing growth and savings.

We test the empirical implications of the model and, more importantly, assess the contribution of
intangible and tangible investment ratios for the evolution of the $W$-$Y$ ratio using newly constructed data for 21 OECD countries over the period 1860-2015. As predicted by our model, we find that the $W$-$Y$ ratio is positively affected by both tangible and intangible investment ratios and negatively affected by income growth. These findings are robust regardless of estimation period, choice of estimator, instrumentation, and inclusion of control variables in the regressions. Furthermore, annual coefficient estimates reveal that the coefficients of the investment ratios move within a very narrow band around the mean value over time. For intangibles, this suggests that even if intangibles were a small fraction of GDP in the 19th century, their identifying variations were sufficiently important to influence the evolution of the $W$-$Y$ ratio, even back then.

Quantifying the contributions of investment ratios to the increasing wealth inequality in the post-WWII period, we find that the surge in investment in intangibles has contributed to a 53% increase in the $W$-$Y$ ratio, while the decline in the tangible investment ratio since 1964 (1980) has resulted in a 26% (19%) decline in the $W$-$Y$ ratio. Thus, the marked expansion of the knowledge economy has overtaken tangibles as a key driver of the $W$-$Y$ ratio.

*University of Western Australia*
*University of Bologna*
*University of Perugia*
References


