Consumption Inequality in the Digital Age*

Kai Arvai† — Banque de France

Katja Mann‡ — Copenhagen Business School

15 August 2023

Abstract

This paper studies how digitalization affects consumption inequality. While previous literature has documented that technological change leads to U-shaped income polarization, we show that this pattern does not translate into consumption or welfare. This is due to price changes that are more beneficial for richer households. By assembling a novel dataset of digital technology in consumption, we establish several new stylized facts: First, high-income households consume a larger share of digitally produced products. Second, consumption items that rely more on digital inputs witness lower price inflation. Third, high-income households spend a larger share of their time on digital-intensive activities. Building on these findings, we present a structural model that quantifies the impact of digitalization on consumption inequality. The model weighs U-shaped income polarization against inflation rates that decrease with income. As a result, the welfare response to digitalization is J-shaped.

Keywords: Digitalization, Inequality, Consumption, Automation, Inflation

JEL Classification: E21, E22, J31, O33, O41

*We would like to thank Melanie Arntz, Christian Bayer, Gabriele Guaitoli, Yannick Kalantzis, Keith Kuester, Moritz Kuhn, Agnieszka Markiewicz, Morten Olsen, Klaus Prettner, Pontus Rendahl and Christoffer Weissert for helpful comments and suggestions. We are grateful to Xavier Jaravel and Kirill Borusyak for sharing their CEX and BLS concordances and to Joachim Hubmer for sharing his production dataset. A big thanks also goes to Pedro Salgado for excellent research assistance. The views in this paper are the ones of the authors and do not represent the views of the Banque de France.

†Banque de France, E-mail address: kai.arvai@banque-france.fr. I gratefully acknowledge funding and support from the DFG Research Training Group "The Macroeconomics of Inequality" during the time of my dissertation.

‡Copenhagen Business School, E-mail address: kma.eco@cbs.dk

1 Introduction

Digital technology transforms our economy, as it fundamentally changes the way we produce and consume. The increased usage of computers and other digital assets in production has resulted in wage and employment polarization in the Western world (Michaels et al., 2014; Akerman et al., 2015; Gaggl and Wright, 2017; Burstein et al., 2019).[1] Over the last decades, the top and the bottom of the income distribution have gained in terms of income and employment shares, whereas routine-intensive jobs in the middle of the distribution have declined. However, progress in digital technology may not only have an *income effect*, but also a *price effect*: Sectors that rely more on digital assets can potentially produce at a lower cost and offer lower prices to their customers. As consumption patterns differ across the income distribution, this will affect richer and poorer households differently and could either reinforce or counteract income changes. This paper studies how the increased usage of digital technology in the production process affects consumption inequality and welfare in the United States.

By combining household- and sector-level data, we establish that high-income households benefit more from price changes, as they consume goods and services that are produced with a higher intensity of digital capital. The bottom third of the income distribution is particularly negatively affected as their consumption bundles feature a low digital intensity and larger price increases. To assess the importance of the price effect for consumption and welfare, in particular compared to the income effect, we present a dynamic general equilibrium model with heterogeneous workers, which we calibrate to the US economy between 1960 and 2017. When taking income-specific price changes into account, the well-known U-shape of income changes translates into a J-shape for welfare: At the top, income gains from digitalization get amplified through favorable price dynamics, whereas they get muted at the bottom.

A priori it is unclear how price changes induced by digitalization impact consumption inequality. As the increased use of digital technology makes some consumption goods cheaper than others, it will benefit the income groups that consume relatively more of these goods. Depending on what those goods are, either rich or poor households could be the beneficiaries. Identifying the direction of the price effect is therefore first and foremost an empirical task, which requires a measure of the digital content of goods and services. Using data from the U.S. Bureau of Economic Analysis (BEA), we identify capital goods that relate to Information and Communication Technology (ICT) and compute their share in industry-level capital stocks. We document that the share of ICT capital in the overall capital stock has increased substantially between 1960 and 2017. Importantly, there is large heterogeneity in the usage of ICT capital across industries. Most services, the information industry and the computer manufacturing industry are ICT-intensive, whereas other manufacturing industries and agriculture are less digitalized. We account for the digital content of intermediate products by leveraging the input-output structure of the production network. The resulting measure covers about 400 final commodities.

In the next step, we link our dataset to consumption categories in the Consumer Expenditure

[1]When considering automation more broadly, the empirical evidence is even ampler, see e.g. Goos and Manning (2007), Autor et al. (2008), Goos et al. (2009) and Autor and Dorn (2013). Prettner and Strulik (2020) and Moll et al. (2022) show that automation also increases wealth inequality. Digitalization is not equivalent to, but a large subset of automation. About two thirds of new automation patents relate to digital technology (Mann and Püttmann, 2023).

Survey (CEX). Using CEX data on household income for 1996-2017, we construct the overall ICT share of consumption baskets along the income distribution. We find that rich households have a 13% larger ICT share in consumption than poor households, with particularly large differences between the bottom 30% and the rest. We also document that consumer price inflation has been weaker for ICT-intensive commodities. A back-of-the-envelope calculation reveals that the price indexes of different income groups diverge by 3 percent over 20 years. We provide an empirical estimate of the price effect via a compensatory variation that computes how much additional income households need to receive to be indifferent to price increases. The poorest households need to be compensated by about 80% of their initial income, whereas the rich require only 40%.

Expenditure data do not allow us to capture the consumption of goods and services without a market price, such as freely available applications or software. In order to take these categories into account, we repeat the empirical exercise using data from the American Time Use Survey (ATUS). There is no evidence that poor households spent a systematically larger share of their time with these kinds of activities than rich households. In contrast, we find evidence that richer households spend a larger fraction of their time on digitally-intensive activities such as household management and administrative volunteering while the poor use more time on low-ICT intensive leisure activities.

In the second part of the paper, we present a structural model to assess the welfare effect of digitalization for different income groups. While the data allow us to get an estimate of the price effect only, the model generates changes in both prices and incomes. The model features a two-sector economy with two types of capital, ICT and non-ICT. Sector 2 uses ICT capital more intensively. The economy is populated by two types of households, high-skill and low-skill, where low-skill households are heterogeneous and sort endogenously into routine and non-routine manual occupations. Routine labor and ICT capital produce a composite good, which is subsequently combined with high-skill and manual labor. The resulting composite is then used together with non-ICT capital to produce the final good. In this set-up, we follow Jaimovich et al. (2021). Digitalization is modeled as an increase in the rate of transformation of output into ICT capital, which affects on the one hand the factor demands and on the other hand the relative price of good 1. We introduce non-homothetic preferences as in Boppart (2014), which imply that the effect of changing relative prices depends on the household's position in the income distribution.

We calibrate the model to the U.S. economy between 1960 and 2017. We use the method of simulated moments to match key moments in the data, in particular the increase in the skill premium, the decline in the price of ICT-intensive goods and differences in consumption between income groups. The model generates income polarization and heterogeneous inflation dynamics that are very close to the data.

While the income effect is about as large for households at the top and the bottom of the distribution, the price effect differs strongly. The poorer the household, the higher the increase in costs of living. As a result, the overall welfare effects over the simulation period vary dramatically. High-skill households experience an increase in welfare that is equivalent to 87% of their initial income, whereas welfare of households at the bottom increases only by about 50%. The middle of the income distribution is the worst off, experiencing weak income growth jointly with high price increases. Their welfare increases only by about 10%. These results imply that the price effect redefines the winners and losers from digitalization, resulting in a J-shape for welfare gains.

We make two contributions to the literature. First, we provide a comprehensive assessment of the welfare effects of automation. Many papers have documented the unequal gains from automation across the income distribution, focusing on changes in employment and wages (Goos and Manning, 2007; Goos et al., 2009; Autor et al., 2008; Autor and Dorn, 2013; Michaels et al., 2014). We relate most closely to two papers: Eden and Gaggl (2018), who study the welfare gains of digitalization in a representative agent framework where technology leads to a reallocation of income across production factors, and Jaimovich et al. (2021), who construct a heterogeneous agents model to consider the effect of automation for welfare of different worker groups.[2] We introduce price effects as a second channel generating heterogeneous welfare effects by adding multiple consumption goods and non-homothetic preferences. We show that this channel is sizable and that it redefines the winners and losers from automation. Hubmer (2023) chooses a similar approach by considering income and expenditure side in a study of the long-run behavior of the labor share.

Second, we identify the technology content of consumption goods and establish digital technology as a driver of inflation and consumption inequality. Several recent papers have documented that inflation rates are higher for low-income than for high-income households (Kaplan and Schulhofer-Wohl, 2017; Jaravel, 2019; Argente and Lee, 2021). Less is known about what explains these differences. Candidate explanations involve the direction of innovation (Jaravel, 2019) and trade (Borusyak and Jaravel, 2021). In studying the heterogeneous effect of digitalization on consumption, we also contribute to the literature documenting and explaining differences in consumption patterns between rich and poor households (Krueger and Perri, 2006; Aguiar and Bils, 2015; Attanasio and Pistaferri, 2014; Meyer and Sullivan, 2023). To our knowledge, we are the first to relate consumption inequality to technological trends, both theoretically and empirically. We also highlight the importance of distinguishing between nominal and real consumption inequality.

In what follows, Section 2 presents the empirical analysis. In Section 3 we introduce our model. Section 4 explains the calibration, followed by the simulation results in Section 5. Section 6 concludes.

2 Empirical Analysis

We assemble a novel dataset to study the share of digitally-produced goods in the consumption baskets of households. Our approach combines industry-level capital stocks with household-level consumption and proceeds along the following steps: First, we create an industry-level measure of ICT intensity by computing the share of ICT capital in an industry's total capital stock. Then, we trace linkages across industries to create an ICT intensity at the level of final commodities. Finally, we match commodities to consumption good categories and calculate the digital share of consumption baskets along the income distribution. Our data reveal that rich households consume more ICT-intensive goods than poor households. At the same time, prices of ICT-intensive goods have grown at a slower pace than prices of non-ICT goods, making the rich households the

[2] See also Vom Lehn (2020) for a similar set-up of the production function. An earlier reference is Krusell et al. (2000).

prime beneficiaries of digitalization. We use a compensatory variation to provide a first estimate of the welfare effects of digitalization along the income distribution. To cover consumption categories without market price, such as freely available goods or home production, we link our digitalization measure to time-use data. We document that richer households tend to spend more time on digital intensive activities, confirming the pattern based on expenditure data.

2.1 Digitalization at the Industry Level

The BEA provides data on the stock of 96 different types of capital for 61 private industries in the Detailed Data for Fixed Assets. Capital stocks are constructed from investment data via careful application of the perpetual inventory method, taking into account quality changes and the introduction of new products.[3] We focus on equipment capital and use annual data between 1960-2017 to construct the stock of digital capital for each industry. We define **ICT capital** to be the following capital categories: Mainframes, PCs, DASDs, printers, terminals, tape drives, storage devices, system integrators and intellectual property products, such as: prepackaged software, custom software, own-account software, semiconductor and other component manufacturing, computers and peripheral equipment manufacturing, other computer and electronic manufacturing, n.e.c., software publishers and computer systems design and related services. We also refer to these assets as digital capital. All other assets are defined as non-ICT capital. A similar approach is chosen by Eden and Gaggl (2018) and Aghion et al. (2020).[4]

Panel (a) of Figure 1 plots the aggregate stock of ICT and non-ICT capital in the U.S. economy by year, indexed to 1 in 1995. Over the whole time horizon, ICT capital has grown much faster than non-ICT capital. Since 1995, the ICT capital stock has increased by more than 700%, whereas non-ICT capital has increased by around 300%.

We measure the degree of digitalization as the share of ICT capital in the overall capital stock of an industry or the whole economy, referring to this measure as **ICT intensity**. It captures how important digital capital is to industries in their production processes. Our measure has the advantage that all data series are directly observable as firms usually own their capital rather than renting it.[5] Since both types of capital are evaluated at their current prices, the measure describes how much ICT capital is worth to producers relative to non-ICT capital. The relative valuation reflects how productive each capital type is, i.e. the level of ICT- and non-ICT technology.

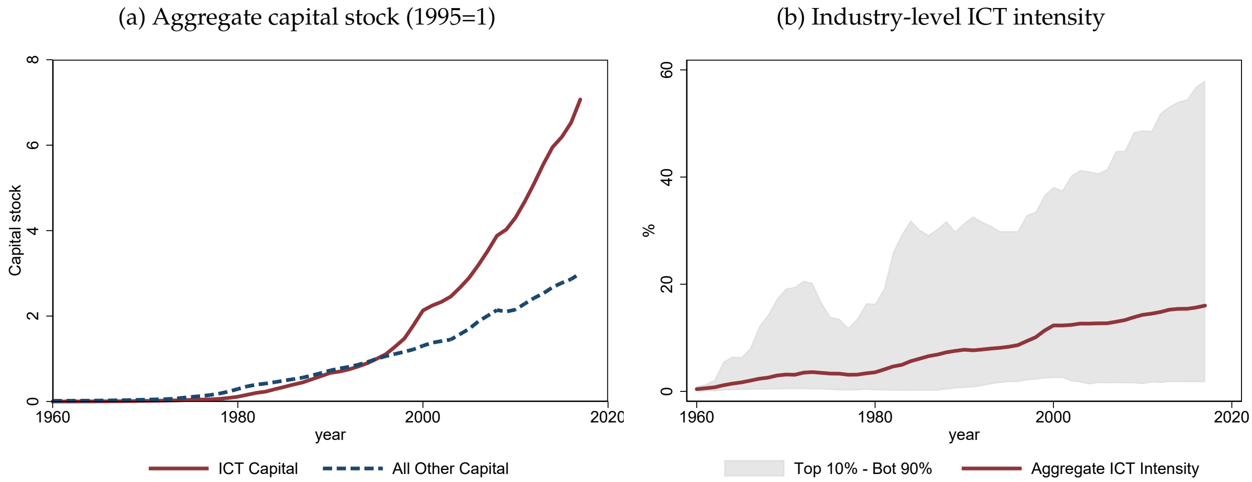
The right panel of Figure 1 shows the ICT intensity over time. We show the aggregate ICT intensity (red solid line) as well as industries at the 10th and 90th percentile of the distribution (gray area). The average ICT capital share has risen substantially from almost zero in 1960 to 16% in 2017.

[3] Detailed information on the methodology underlying the construction of these estimates is available at https://www.bea.gov/system/files/methodologies/Fixed-Assets-1925-97.pdf. Adjustments in quality are particularly relevant in the case of digital capital, which the BEA methodology takes into account.

[4] Alternatively, goods could be classified based on whether they themselves constitute digital goods. Appendix A.4 discusses this issue in more detail. The input-based measure is better suited to compare how changes in production technology affects incomes and prices.

[5] This is not the case when focusing on the ICT capital income share in value added, which has been used in some papers to measure digitalization, see e.g. Eden and Gaggl (2018) and Karabarbounis and Neiman (2019). As there is no market to directly observe the rental rate of production capital, these papers need to estimate returns. Depending on the underlying methodology and data, estimates differ widely across papers.

Figure 1: ICT capital



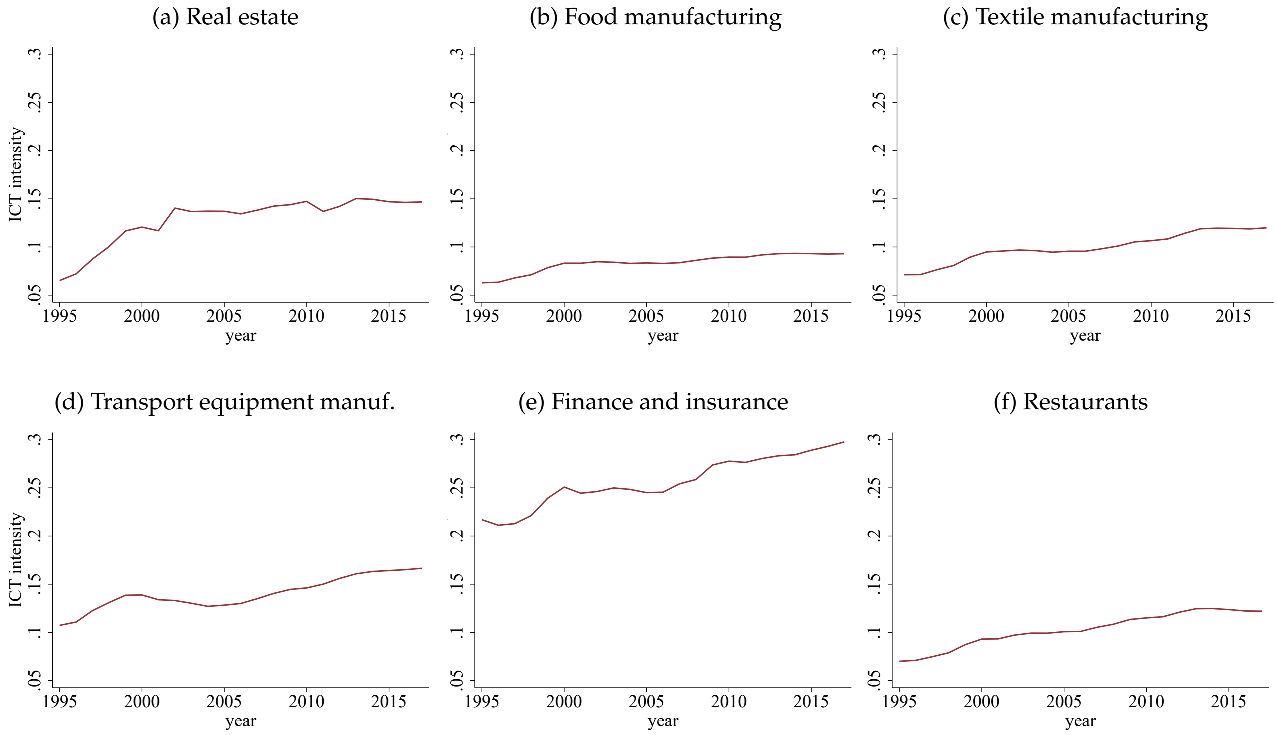
Note: The left graph shows ICT and non-ICT capital in the U.S. economy between 1960 and 2017. Both series are normalized to 1 in 1995. The right graph shows the share of ICT capital in the total capital stock (ICT intensity) by BEA industry. The solid line shows the average and the gray area the industries between the 10th and 90th percentile of the distribution. *Source:* BEA and own calculations.

Underlying the aggregate measures is a large degree of heterogeneity across industries. Some industries have barely accumulated any ICT capital, while in other industries, more than half of the capital stock consists of ICT capital by 2017.[6] Industries in the top 10% of the ICT intensity mostly belong to the finance and insurance industry or to the computer and electronic products manufacturing industry. Among the least ICT-intensive industries are agriculture, the plastic and rubber industry and the textile industry. In Appendix A.2, we compare our ICT intensity measure to two other established measures of automation. We show that there is a positive correlation with the patent measure of Mann and Püttmann (2023) and with the share of cognitive tasks in an industry. We also discuss how our measure differs from other capital intensities used in the literature, as for example by Hubmer (2023) or Aghion et al. (2020).

Industries may not only use digital capital in their own production activities, but also use intermediate inputs that have been produced with digital capital. In order to calculate the share of digital capital used in the production of final commodities, we need to take input-output linkages among industries into account. The BEA's Input Output (IO) Accounts show how industries provide input to or use output from each other. We use the detailed IO tables after re-definitions. The tables are published every five years and contain around 400 private industries, which can be matched to the 61 BEA industries[7]. We create a commodity-by-commodity direct requirements matrix, which we use iteratively to calculate the digitalization share of final goods and services as a weighted mean of the digitalization shares of its intermediate inputs and value added. Commodities that initially have a low degree of digitalization tend to be more digitalized after the inputs from other industries are considered. The reverse is true for highly digitalized industries. Appendix A.1 provides further details on the dataset and our procedure.

[6]The temporary drop for the top 10 % in the 1970s is due to declining ICT investment activity and high depreciation rates in the sector for credit intermediation and in scientific research.
[7]All sectors related to housing are matched to the real estate sector. For 1996-2004 we use the 1997 IO table for 2005-2017 we use the 2012 IO table.

Figure 2: ICT intensity by commodity



Note: The graph shows ICT intensities for six broad commodity categories over time. The ICT intensity in these categories is calculated as a weighted average of ICT intensity of the relevant IO commodities. The weights correspond to their share in value added. *Source:* BEA and own calculations.

We construct the ICT-intensity measure for all IO commodities between 1996 and 2017. This time period is chosen because we will later link the industry data to the CEX, which is only available for more recent years. We group commodities into 25 broader categories and plot the ICT intensity of six of them over time in Figure 2. In anticipation of the link to consumption categories in the CEX in the next section, we consider the six most important categories for consumption: Real estate, food manufacturing, textile manufacturing, transport, finance and insurance and restaurants. There has been an increase in the ICT intensity in all of the commodity categories over time and the pattern often looks similar, e.g. reflecting the build-up and subsequent burst of the dotcom bubble. However, commodities are characterized by large differences in the average ICT intensity. In particular finance and insurance has a high ICT intensity throughout the whole sample period, whereas food manufacturing is characterized by a low ICT intensity. These differences will become relevant in the construction of ICT intensities for individual consumption baskets.

2.2 Digitalization and Consumption

We match final commodities to consumption categories to calculate the digitalization share of consumption along the income distribution. We measure expenditures of U.S. households using the Consumer Expenditure Survey (CEX). The CEX is the most detailed expenditure survey in the United States, carried out at the household level. Next to data about the purchases of hundreds of disaggregate goods and services, it also contains a large amount of demographic and income

Table 1: Expenditure shares for 1st and 10th decile by commodity category in %

industry	1997			2017		
	1st	10th	**ratio**	1st	10th	**ratio**
real estate	21.49	24.76	**0.87**	26.06	28.76	**0.91**
food/beverages/tobacco	17.69	8.75	**2.02**	13.98	7.21	**1.94**
transport equipment	7.59	9.29	**0.82**	6.83	7.95	**0.86**
restaurants	7.33	7.87	**0.93**	7.78	7.98	**0.98**
textile/apparel/leather	7.30	5.78	**1.26**	5.07	3.90	**1.30**
chemicals/petroleum	6.31	4.86	**1.30**	7.23	4.55	**1.59**
finance/insurance	5.16	7.26	**0.71**	5.49	6.98	**0.79**
utilities	4.86	2.97	**1.64**	5.52	2.69	**2.05**
information	4.35	3.70	**1.18**	5.67	3.93	**1.44**
misc. manufacturing	3.02	3.87	**0.78**	1.88	2.10	**0.90**
other services	2.94	4.28	**0.69**	3.17	5.03	**0.63**
machinery/electrics/electronics	2.42	2.82	**0.86**	1.94	2.64	**0.74**
agriculture	2.31	1.22	**1.89**	2.31	1.35	**1.72**
health	1.77	1.71	**1.03**	1.47	2.42	**0.61**
trade/transport/warehousing	0.93	1.70	**0.55**	0.94	2.02	**0.46**
wood/furniture	0.87	1.63	**0.53**	0.92	1.23	**0.75**
professional services	0.79	0.67	**1.18**	0.26	1.02	**0.25**
paper/printing	0.64	0.48	**1.32**	0.58	0.35	**1.65**
education	0.60	1.36	**0.44**	0.53	3.11	**0.17**
rental and leasing	0.58	1.51	**0.39**	0.58	1.03	**0.57**
rubber/nonmetallic minerals	0.43	1.13	**0.38**	0.61	0.46	**1.35**
admin support	0.24	0.47	**0.50**	0.39	0.60	**0.66**
arts/entertainment/recreation	0.21	1.01	**0.21**	0.48	1.51	**0.32**
accommodation	0.16	0.86	**0.19**	0.30	1.16	**0.26**
primary and fabricated metal	0.02	0.04	**0.45**	0.03	0.05	**0.61**

Note: Values for each year represent three-year averages (e.g. for 1997 we use 1996-1998). Commodities are ranked by their importance to low-income households. *Source:* CEX.

information on the households. This enables us to study the expenditure pattern of households by income. We use the public-use microdata for 1996-2017. We combine the interview survey, which contains information across all expenditure categories, with the diary survey, in which households report purchases on only a subset of goods and services, but in a more detailed way. There are around 2,500-3,000 households per year in each of the surveys.

As explained in more detail in Appendix A.1, we divide households into equal-sized bins based on total household income. For each income group, we create a weighted average of expenditure by year. Using the concordance table of Borusyak and Jaravel (2021), we link 809 CEX commodities to 159 IO commodities. Each commodity in the CEX is matched to a unique IO industry.[8]

Income groups vary by their spending pattern. Table 1 shows for 25 broad commodity categories the expenditure share of households in the first and in the tenth decile of the income distribution as well as their ratio, both for the average around 1997 and for 2017. The largest share of income is being spent on housing and therefore gets allocated to real estate. There has been some increase

[8]Note that the concordance table treats housing as a service from the real estate sector rather than as a good from the construction sector. Housing is also special since it serves both as consumption and investment good for house owners. We take this into account by using the rental value of owner-occupied housing, which equals its consumption value.

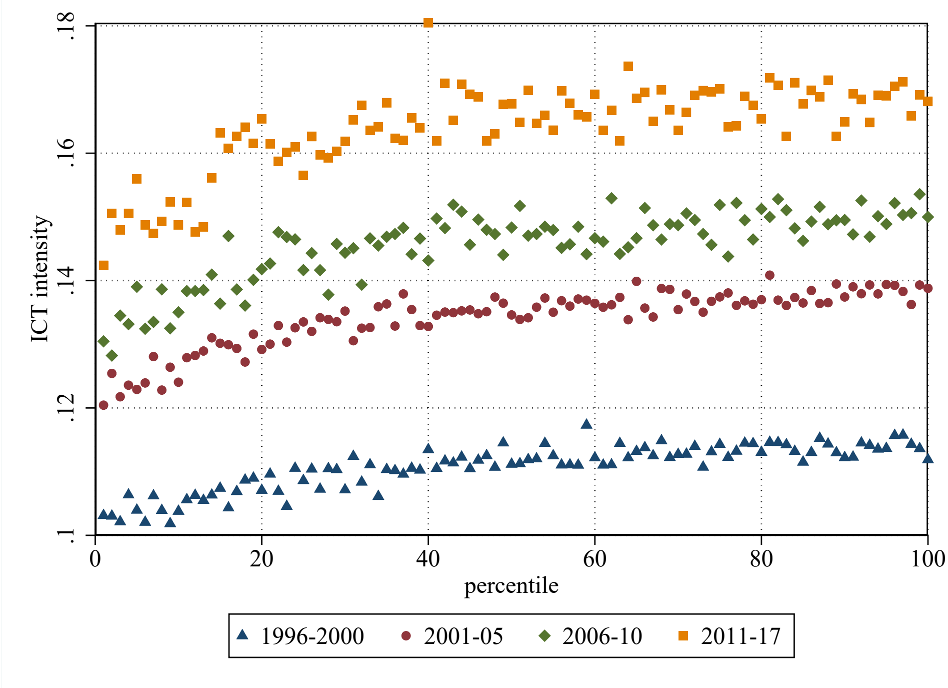
Figure 3: ICT intensity vs. relative expenditure shares by commodity in 2017

Note: Each circle corresponds to one of the 159 IO commodities. The size of the circle reflects the share of expenditures for that category of the median household. The x-axis shows the log ratio of the expenditure share of the top 10 % relative to the bottom 10 %. The more to the right the data are, the more important that category for the rich. Values around zero indicate that a category is equally important to the top 10% as it as for the bottom 10 %. The y-axis shows ICT intensity of the industry, the horizontal green line indicates average ICT intensity. Data are for 2017. *Source:* BEA, CEX and own calculations.

in importance of this category over time, whereas the ratio of the two deciles has been stable. Other important categories with similar spending patterns across deciles are transport equipment – which most notably includes expenditure on cars –, restaurants as well as information – comprising of telecommunication, IT services, broadcasting, audio and video recordings. Some categories are consumed more extensively by the rich, in particular accommodation, entertainment and recreation, education and finance and insurance. The consumption basket of the poor is more tilted towards food, beverages and tobacco, utilities and agricultural products. There has been little change in expenditure ratios over time for most categories. Notable exceptions are education, whose share in the consumption basket of the rich more than doubled while roughly constant for the poor, and professional services, whose importance increased for the rich while dropping for the poor.

The ICT intensity of consumption baskets varies systematically across households. Before considering the ICT intensity of different consumption baskets as a whole, we illustrate which commodities are particularly relevant for determining the ICT intensity of high- vs. low-income consumption. In Figure 3, we plot the ICT intensity against the relative importance of a commodity (in terms of expenditure share) for high-income vs. low-income households. We consider data for final commodities in the year 2017. The pattern is very similar for earlier years. Each dot corresponds to an IO commodity and the size reflects the importance of the commodity (in terms of expenditure shares for the median household).

Figure 4: ICT intensity along the income distribution



Note: The graph shows the ICT share of the consumption basket by percentile for different sub-periods.
Source: BEA, CEX and own calculations.

There is a positive correlation between the ICT intensity and the relative expenditure share of high-income households. On the left side of the graph, categories are more important for poor households and most ICT intensities are below the average value of 0.16. These are in particular goods produced by manufacturing industries like textiles or food. Around the center commodities are equally important for both households. There we find industries with very different ICT intensities. The largest category, real estate, has an average ICT intensity, whereas for example the information industry has a very high ICT intensity. Moving to the right part of the graph, which depicts commodities important for the rich, there are some categories of average ICT intensities, but also some with very large ICT intensities such as education or finance and insurance.

As the final step, we compute the ICT intensity of the whole consumption basket by income percentile. This measure reflects the expenditure patterns across income levels and the industry-level ICT intensities. Figure 4 shows the ICT intensity of consumption along the income distribution for four different time periods. At any time, the ICT intensity is substantially lower for the poorest 30% of the households than for richer households. The ICT intensity has increased between 1996 and 2017 across all income percentiles. However, the increase has been slightly larger for higher percentiles. The strong increase in the ICT share over time that we documented in Section 2.1 affects the consumption bundle of the rich more. The inequality in digital consumption has thus been increasing over the last 20 years. There is a 13% difference between top and bottom in 2011-17.

In Appendix A.3 we carry out several robustness checks of our results. First, we exclude durable goods from the construction of the consumption basket. We still see that the ICT intensity increases along the income distribution. This result is largely driven by food expenditures, which feature a low ICT intensity while being disproportionately important for the poor. We also repeat the analysis on a dataset that excludes finance and insurance, which may be considered an investment good as well as a consumption good. Finally, we regress the ICT intensity on income in a panel regression that controls for year fixed effects. The coefficient on ICT intensity is positive and significant, and remains so when additionally controlling for household age. Even though richer households tend to be older and consume more ICT intensive goods and services, income is still a significant predictor of ICT intensity.

2.3 Digitalization and Price Changes

The CEX does not provide separate information on the quantities and prices of the goods purchased. We therefore supplement our dataset using consumer price indices from the BLS.[9] We match CEX product categories with BLS price data series between 1996 and 2017, using the same concordance of Jaravel (2019).[10] We then create Törnqvist price indices for each income percentile. The Törnqvist index is a chain-weighted price index that considers product substitutions and other changes in spending habits, and is therefore well suited for our purposes. The growth rate of the Törnqvist index is a weighted average of the growth rates of the disaggregated price series of I individual goods, p_i, where the weights are nominal expenditure shares of income percentile j, $s_{i,j}$,

$$\pi_{j,t} = \sum_{i=1}^{I} \ln\left(\frac{p_{i,t}}{p_{i,t-1}}\right)\left(\frac{s_{i,j,t} + s_{i,j,t-1}}{2}\right) \tag{1}$$

The level can be recovered recursively relative to a base year 0 with $p_0 = 1$ as $p_{j,t} = p_{j,t-1}\exp(\pi_{j,t})$. Figure 5 (a) shows the evolution of this price index for each income percentile.

Over the 20 years of our sample, most income percentiles experienced an increase in their price index of more than 50%, indicating an average annual inflation rate of slightly more than 2%. There is a negative relationship between a household's position in the income distribution and its individual inflation rate. Poor households experienced faster price increases than rich households. The difference in inflation for the richest and poorest households amounts to around 6 percentage points at the end of the sample period. This is in line with Jaravel (2019) and Kaplan and Schulhofer-Wohl (2017), who find that inflation has been higher for poor households than for rich households in this period.

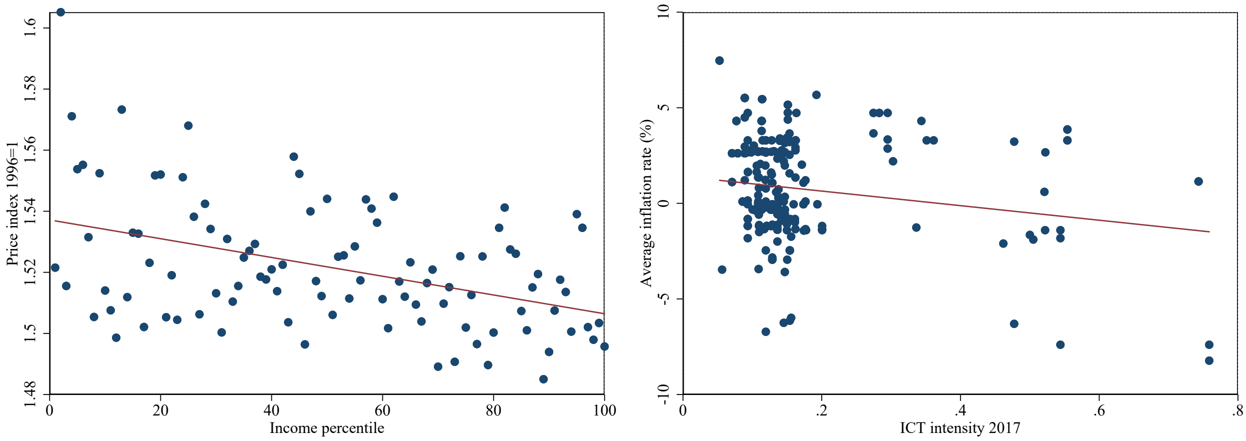
The inflation rates co-move with differences in ICT intensities of the products that households consume. Figure 5 (b) plots the average inflation rate of single items in the CEX against the item's ICT intensity. In a simple regression of inflation on ICT intensity at the commodity level (Ap-

[9] This is a standard procedure, see e.g. Hobijn and Lagakos (2005) and Jaravel (2019), but nevertheless entails certain caveats: The prices in the CEX are those actually paid by the households, whereas those in the BLS are posted prices paid by the average urban consumer. Furthermore, with this procedure we are unable to capture quality changes in goods or the exit or entry of products within narrow BLS categories. Including quality changes would lead to a stronger divergence in the price dynamics of high- and low-ICT goods, which would make the differences in inflation between high- and low-income households even more pronounced. The results presented below should therefore be considered as a lower bound.

[10] Some disaggregated price series start later than 1996. In these cases, we match the CEX category to the series at the next higher hierarchical level in the BLS.

Figure 5: Inflation

(a) Price index in 2017 (1996=1) for income percentiles (b) Inflation 1996-2017 of single items and ICT intensity



Note: The left graph plots the Törnqvist price index in 2017 indexed to 1 in 1996 for each income percentile. The red line shows the prediction from a linear regression. The right graph shows ICT intensity in 2017 against average annual inflation in consumer prices between 1996 and 2017 for single consumption categories. The red line shows the prediction from a linear regression. *Source:* BEA, CEX, BLS and own calculations.

pendix Table A2), the coefficient is negative and significant, indicating that consumption categories with higher ICT intensity experienced lower inflation rates. Our findings align with Aghion et al. (2020), who show that automation reduces the producer price index in French manufacturing industries. Graetz and Michaels (2018) find that the use of robots – another type of automation technology – also leads to lower prices.

Among the categories with the largest price decline are applications, games, computer and electronic products. The largest price increases took place in low-ICT intensity categories such as housing, fuels and gasoline, where obviously, inflation has been driven by factors not related to technological change.

As a robustness check, we also include the import share in Table A2, defined as the value added share of imported goods from the IO use tables. Cheaper imports from countries such as China are an obvious alternative explanation for lower inflation (see e.g. Borusyak and Jaravel, 2021). Indeed, a higher import share is associated with lower price increases. When including the import share in the regression, the coefficient on ICT intensity remains significant but gets closer to zero. Next to including imports as a robustness check, we also consider price data by NIPA at the level of the 61 BEA industries in Figure A8 to confirm the pattern on a more aggregate level. The correlation between ICT intensity and inflation is also negative and significant.

Finally, we conduct a back-of-the-envelope calculation in which we combine the difference in ICT intensities between the bottom and the top income percentile with the slope coefficient of the regression of average price changes on ICT intensity[11]: In a calculation using the correlation

[11] This approach is based on the following reasoning: Both the ICT intensities and the price indices are originally at the level of final commodities. Those are then linked to consumption categories in the CEX which yields ICT intensities of the consumption basket and inflation rates at the percentile level. Differences in consumption ICT intensity are therefore driven by commodity ICT intensities and expenditure shares whereas inflation differences are driven by commodity price changes and expenditure shares. The back-of-the-envelope exercise combines these three variables in one calculation: $(1 + 4.22/100 \cdot 0.03)^{22} = 2.8\%$.

coefficient between price changes and ICT intensity of -4.22 (Table A2), and a difference in ICT intensity in the basket of rich and poor households of 3% (Figure 4), the effect accumulates to a 2.8% difference in the price index from 1996 and 2017. This is almost half of the value in the difference of the actual price index.

2.4 An Empirical Estimate of the Price Effect

This section provides a first estimate of the price effect, asking by how much digitalization makes rich households better off compared to poor households. To this end, we combine our findings on the ICT intensity of consumption in Section 2.2 with the price data of Section 2.3 and carry out two kinds of compensatory variations. In the first one, we fix the composition of the consumption basket for each income percentile at the base year, whereas in the second one, we allow for substitution across expenditure items. In both cases, we depart from the consumption expenditures of each income percentile in 1997. Then, we calculate how much additional income households need to receive in order to be indifferent to the price increases that took place between 1997 and 2017. The first compensatory variation can be approximated by

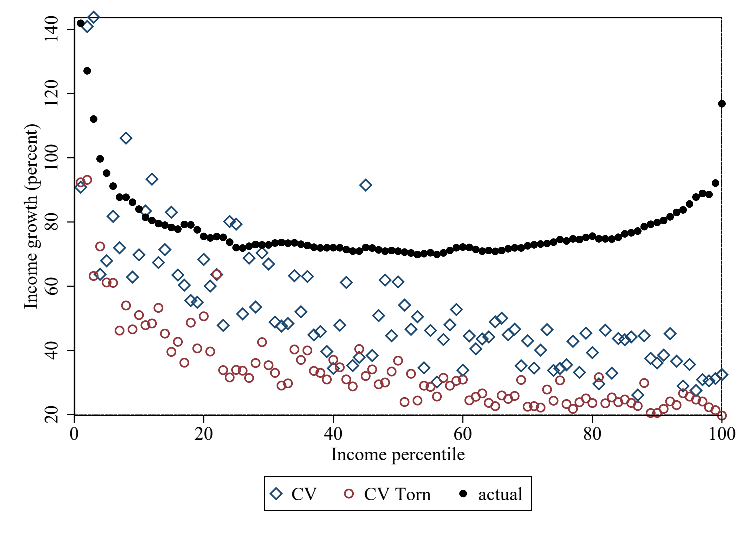
$$CV_j = \sum_i x_{ij} p_i \left(\frac{\Delta p_i}{p_i} \right), \tag{2}$$

where $x_{ij} p_i$ is percentile j's expenditure on commodity i in the base year 1997 (price p_i times quantity x_{ij}). $\frac{\Delta p_i}{p_i}$ is the price change between 1997 and 2017. Appendix A.7 provides a derivation for this formula. According to this equation, the money needed to compensate households for price changes equals the initial consumption bundle times the change in prices for each commodity. For the second compensatory variation that considers substitution behavior, we compute the compensatory income by inflating the expenditures with the income-specific Törnqvist price index.

Figure 6 illustrates the required income growth along the income distribution according to the two compensatory variations, together with the actual nominal income growth in the CEX, which we compute as differences in nominal incomes within percentiles over time. Considering the actual increase in income (black circles), there is clear evidence of a U-shaped polarization. The top and the very bottom of the income distribution experienced substantially larger increases than the middle. This is contrasted by the income required to compensate for price changes.

The compensatory variation without substitution (blue diamonds) demonstrates that price changes have been most disadvantageous for poorer households. The additional income required to buy the same consumption bundle at higher prices decreases with income. The poorest 10 % of households need to be compensated by on average more than 80% of their initial income, which eats up much of their actual income increase (in some cases the values are even above the actual income increase). The richer a household gets, the lower the required compensation for the price increase relative to the initial income. The richest households only need to receive around 40% of their initial income to be compensated for inflation in their basket. The red hollow circles show the additional income required to uphold the real value of expenditures when inflated with the percentile-specific Törnqvist index. Compared to the compensatory variation with fixed expendit-

Figure 6: Compensatory variation and actual income changes



Note: This graph plots the actual increase in nominal income (black circles) for each income percentile between 1996 and 2017, the compensatory variation without substitution (blue diamonds), and the compensatory variation with substitution (red hollow circles). *Source:* BEA, CEX, BLS and own calculations.

ure shares, the required additional income is generally lower, suggesting that substitution plays a role and lowers the costs of higher prices. However, the overall pattern of higher values for the poorer households is still visible.

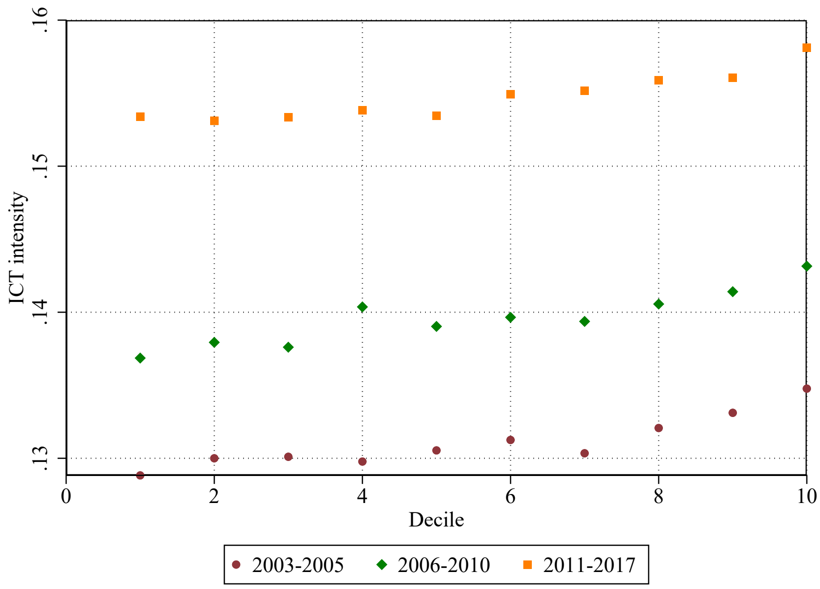
This exercise suggests that the price effect and the different consumption behavior need to be taken into account when assessing welfare effects of digitalization: A U-shaped income polarization has to be weighed against an inflation cost curve that is downward sloping in income. That could make poor households substantially worse off than when only considering income changes. Note, however, that this exercise makes no statement about where income changes originate from; we need the model in Section 3 to establish a link between digitalization and incomes.

2.5 Digitalization and Time Use

So far, we have equated consumption with expenditure. While this is reasonable given the data available to us, it ignores that consumption uses both market expenditures and time as inputs. The opportunity cost of time is lower for individuals working fewer hours (which correlates with income), which is why these individuals are more likely to substitute away from market expenditures (Aguiar and Hurst, 2005). In our setting, the ICT intensity of expenditure and consumption will differ if lower-income households do home production of goods that have a higher- or lower-than-average ICT intensity. For example if poor households are more likely to prepare their own meals and make their own clothes, their expenditure on the low-ICT categories food and textiles are going to be lower, creating an upward-bias of their ICT intensity.

A second worry are goods and services that are available for free, or at a very low marginal cost. These are not properly accounted for in the expenditure data. It can be expected that low-income households consume a larger share of free goods, and if these goods differ in their ICT intensity

Figure 7: ICT intensity for time use along the income distribution



Note: The graph shows the ICT share of the overall time use by income decile for different subperiods between 2003 and 2017 *Source:* BEA, ATUS and own calculations.

from goods that are sold at a market price, it will also bias our measure. In some cases, both a free and a subscription version exist, for example for many apps and software. The lower-quality free version is more likely going to be consumed by low-income households. If the poor spend substantially more time with free digital activities than the rich, their consumption is less adequately represented by an expenditures-based measure.

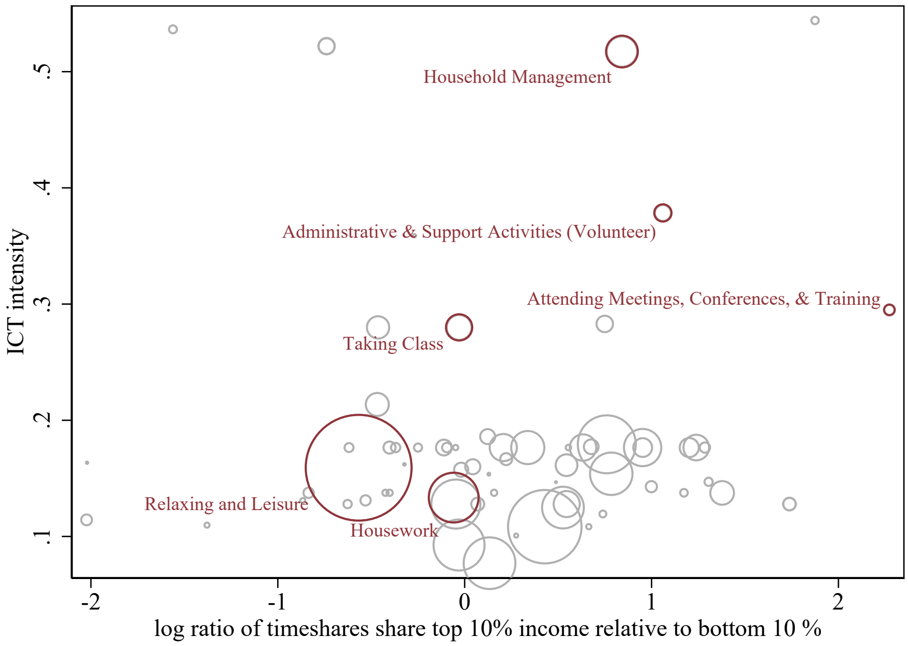
In order to address these concerns, we supplement our analysis using an additional dataset, the American Time Use Survey (ATUS). ATUS documents how Americans allocate their time across different activities. It has been running sine 2003 and is based on a subsample drawn from the Current Population Survey (CPS). Within a 24-hour period, the respondents document their activities in a diary, choosing among more than 400 different activity categories. We replicate the analysis of Section 2.2 by matching the final commodities, for which we have calculated ICT intensities, to the ATUS activity categories. Appendix A.1 provides details on the construction of the dataset. We exclude working, sleeping and government services. While ATUS avoids the measurement error induced when using expenditure data, the disadvantage is a shorter time coverage and – in some cases – less detailed categories.

We compute the average ICT intensity of time use of an individual by summing up the time-share weighted activities multiplied with their ICT intensities. We track the time share of activities since 2003 and compute the average ICT-time-intensity for income deciles[12]. Analogous to the ICT share of expenditures in Figure 4, Figure 7 shows the ICT intensity for time use along the income distribution.

In all three considered subperiods, the ICT time intensity increases along the income distribution, just as the ICT *expenditure* intensity. While the ICT *expenditure* intensity in Figure 4 increases only

[12]The household income data of ATUS are reported in few income bins, which is why we consider deciles here.

Figure 8: ICT intensity vs. relative time shares by 4 digit activity-code in 2017



Note: Each circle corresponds to an aggregate 4-digit ATUS category. The size of the circle reflects the share of overall time use for that category of the median household. The x-axis shows the log ratio of the time use share of the top 10 % richest relative to the bottom 10 %. The more to the right the data are, the more important that category for the rich. Values around zero indicate that a category is equally important to the top 10% as it as for the bottom 10 %. The y-axis shows ICT intensity of the category. Data are for 2017. *Source:* BEA, ATUS and own calculations.

for the first three deciles, ICT *time use* intensity increases along the whole income spectrum. The difference between the poorest and richest decile is around 0.5 percentage points at any point in time, which is smaller than the difference in ICT expenditure intensities. Nevertheless, the coefficient on income is significant in a regression shown in Appendix Table A3. When additionally controlling for age, the correlation between income and ICT time intensity becomes stronger, as old households tend to pursue activities that are less ICT-intensive. Our findings are also robust to only considering employed respondents.

It is reassuring that the results based on time use are in line with those based on expenditure data. This implies that even though we are disregarding some parts of consumption when focusing on expenditure, we still get a fairly accurate picture of the differences in ICT intensities of the consumption baskets of rich and poor households. In Section 3, we will therefore for simplicity assume that all consumption shows up in form of purchases of goods or services at a market price.

Nevertheless, it is informative to uncover which activities are driving the results of Figure 7. Figure 8 plots the ICT intensity of 4-digit activity codes against their relative importance for the rich vs the poor, analogous to Figure 3. Some categories are more important for the rich and feature at the same time high ICT time intensities. Among those categories are attending conferences and training, administrative and support activities in volunteering (including computer use) and household management (including financial management and e-mail messages). The category with the largest time use by the median household is relaxing and leisure, which is slightly more important for poorer households. This category features some high-ICT intensity activities such as computer use for leisure or playing games, but also activities such as reading, arts and crafts, or tobacco and drugs use, which have a low ICT intensity. Furthermore, it contains television and movies, which has an average ICT intensity according to our measure, owing to the fact that the film industry uses many low-ICT inputs.

Our results also speak to the importance of goods and services that are available for free. While ATUS does not provide information on whether the activities include a purchase or not, it can be assumed that several categories have low marginal costs, such as computer use, e-mail services, playing games. These high-ICT activities are not systematically more important for either rich or poor households, which is why we do not expect them to bias our results. Housework – another activity without market price – features a low ICT intensity and tends to be more important for poor households. It therefore creates an upward bias of the ICT intensity of poor households, which means that the ICT intensity increases more strongly along the income distribution than the expenditure-based measure suggests. The results presented in Section 2.2 thus constitute a lower bound.

By focusing on activities, this section has tried to provide a more inclusive measure of consumption and how it is affected by digitalization. However, multiple additional dimensions remain along which digitalization could affect consumption inequality. These include the provision of public services, the availability of information, or transaction costs such as time spent shopping, waiting or on job search. While Appendix A.6 touches upon this last point, it is beyond the scope of this paper to comprehensively assess all of these different dimensions, and we leave further exploration for future research.

3 Model

In the following, we model the effect of an increase in digital technology on consumption inequality, working through changes in relative prices and income. In the model economy, there are two capital goods (ICT and non-ICT) and two sectors with different factor inputs. Each sector produces a unique final consumption good. Households have non-homothetic preferences over the two goods, which means that income differences induce differences in consumption patterns. There are two types of households, high-skill and low-skill, where low-skill households sort into routine and manual occupations depending on ability. Technological advances affect the demand for factor inputs and household incomes. This setting allows us to quantify the effect of digitalization on consumption inequality and to analyze the mechanisms at work.

3.1 Sectoral production functions

Two sectors $i = 1, 2$ each produce a final output using as inputs two types of capital and three types of labor: (Low-skill) routine labor R_i, (low-skill) manual labor M_i and (high-skill) cognitive labor H_i. Production is nested.

Inner nest: Routine labor R_i and ICT capital ICT_i are used to produce an automated work output

(AW):

$$AW_i = \left[\phi_i R_i^{\frac{\eta_i - 1}{\eta_i}} + (1 - \phi_i) ICT_i^{\frac{\eta_i - 1}{\eta_i}} \right]^{\frac{\eta_i}{\eta_i - 1}}, \tag{3}$$

with $\phi_i \in (0,1)$ weights that govern the importance of each input factor, and η_i the elasticity of substitution between routine labor and ICT. If η_i is large, then routine labor tends to be more substitutable to ICT. If $\eta_i = 1$, the production function is Cobb-Douglas with factor share ϕ_i.

Middle nest 1: High-skill labor H_i and the automated work output are used to produce an intermediate product SW:

$$SW_i = \left[\gamma_i H_i^{\frac{\epsilon_i - 1}{\epsilon_i}} + (1 - \gamma_i) AW_i^{\frac{\epsilon_i - 1}{\epsilon_i}} \right]^{\frac{\epsilon_i}{\epsilon_i - 1}}, \tag{4}$$

with weights $\gamma_i \in (0,1)$ and ϵ_i the elasticity of substitution between high-skill labor and AW. Smaller values indicate that both inputs are complements, larger values that they are substitutes

Middle nest 2: SW and manual labor are combined to produce total work output (TW) in a Cobb-Douglas manner with factor share ψ_i:

$$TW_i = M_i^{\psi_i} SW_i^{1-\psi_i}$$

which is motivated by the observation of Eden and Gaggl (2018) and Jaimovich et al. (2021) that the share of non-routine manual labor in national income has stayed fairly constant over time.

Outer nest: Total work and non-ICT capital are combined in a Cobb-Douglas production function, owing to the fact that the share of non-ICT capital in revenue has been constant over the last decades (Koh et al., 2020):

$$Y_i = K_i^{\alpha_i} TW_i^{1-\alpha_i} \tag{5}$$

There is perfect competition in each sector. Firms demand a price P_i for their good, where we normalize $P_2 = 1$ without loss of generality. They pay wages W_H, W_R and W_M and rent capital which pays interest rates R_K and R_{ICT}. First-order conditions yield the following equations for factor returns

$$K_i: \quad P_i \alpha_i \frac{Y_i}{K_i} = R_K \tag{6}$$

$$M_i \quad P_i (1 - \alpha_i) \psi_i \frac{Y_i}{M_i} = W_M \tag{7}$$

$$H_i: \quad P_i \gamma_i (1 - \alpha_i)(1 - \psi_i) \frac{Y_i}{SW_i^{\frac{\epsilon_i - 1}{\epsilon_i}} H_i^{\frac{1}{\epsilon_i}}} = W_H \tag{8}$$

$$R_i: \quad \frac{P_i \cdot AW_i^{\frac{\epsilon_i - 1}{\epsilon_i}}(1-\gamma_i)(1-\psi_i)(1-\alpha_i)\phi_i Y_i}{R_i^{\frac{1}{\eta_i}} AW_i^{\frac{\eta_i - 1}{\eta_i}} SW_i^{\frac{\epsilon_i - 1}{\epsilon_i}}} = W_R \tag{9}$$

$$ICT_i: \frac{P_i \cdot AW_i^{\frac{\epsilon_i - 1}{\epsilon_i}}(1-\gamma_i)(1-\psi_i)(1-\alpha_i)(1-\phi_i) Y_i}{ICT_i^{\frac{1}{\eta_i}} AW_i^{\frac{\eta_i - 1}{\eta_i}} SW_i^{\frac{\epsilon_i - 1}{\epsilon_i}}} = R_{ICT} \tag{10}$$

3.2 Technological progress and capital formation

The output of sector 1 can be used for consumption as well as to produce non-ICT capital while the output of sector 2 can be used for consumption and for the production of ICT capital. While we assume that Y_1 can be transformed at the same rate (of 1) into the investment good and the consumption good, we define μ as the rate of transformation of Y_2 into ICT capital. Then, the resource constraints are

$$Y_1 = C_1 + I_K, \qquad Y_2 = C_2 + \mu I_{ICT}. \tag{11}$$

μ is the relative price of ICT capital and at the same time measures exogenous progress in ICT technology (see e.g. Greenwood et al., 1997; Karabarbounis and Neiman, 2014; Eden and Gaggl, 2018; Cortes et al., 2017 for similar set-ups). A decline in μ will make ICT technology more productive. We do not consider any other sources of growth in the economy, such as TFP growth or sector-specific productivity growth.[13]

Capital follows the standard law of motion

$$K' = (1-\delta_K)K + I_K, \qquad ICT' = (1-\delta_{ICT})ICT + I_{ICT}, \tag{12}$$

where δ_K and δ_{ICT} are the depreciation rates.

3.3 Households

Occupational Choice

There are two types of households in this economy, high-skill and low-skill, which are of fixed (but time-varying) sizes $\bar{H}$ and $\bar{L}$. We think about skill in terms of level of education, which is determined before entering the labor market. Households can work in any sector and can switch sectors at no cost. Following Jaimovich et al. (2021), we assume that high-skill households are all identical and always allocate to the high-skill tasks, whereas low-skill households may work in

[13]In our model, TFP growth for example would lead to an increase in overall income. Together with non-homothetic preferences, a general increase in the consumption share of ICT goods would follow, which would lead to an increase in the price of ICT-intensive goods. This is not in line with the data. Sector specific growth could explain the decline of ICT consumption prices, but cannot explain why the price of ICT investment has declined even more.

either routine or manual occupations. Each low-skill household j is endowed with two idiosyncratic productivity parameters $\lambda_{R,j}$,$\lambda_{M,j}$ indicating their routine and manual abilities, which are drawn from a joint distribution $\Gamma(\lambda_R, \lambda_M)$. More able households supply more units of labor and therefore earn a higher wage. A household chooses to work in R if

$$\lambda_{R,j} W_R \leq \lambda_{M,j} W_M \tag{13}$$

In consequence, total supply of routine labor (in efficiency units) is

$$\bar{R} = \bar{L} \int_{-\infty}^{\infty} \int_{-\infty}^{\lambda_M(\lambda_R)} \lambda_R \Gamma'(\lambda_R, \lambda_M) d\lambda_M d\lambda_R \tag{14}$$

where $\lambda_M(\lambda_R)$ denotes the cutoff in ability for each λ_R value such that, below it, workers choose to work in R and not in M. $\Gamma'(\lambda_R, \lambda_M)$ is the cumulative distribution function corresponding to the probability distribution function $\Gamma(\lambda_R, \lambda_M)$.

Correspondingly, the total number of manual efficiency units provided is

$$\bar{M} = \bar{L} \int_{-\infty}^{\infty} \int_{-\infty}^{\lambda_R(\lambda_M)} \lambda_M \Gamma'(\lambda_R, \lambda_M) d\lambda_R d\lambda_M \tag{15}$$

with $\lambda_R(\lambda_M)$ defined accordingly.

Preferences

All households have non-homothetic preferences over the two goods 1 and 2. We follow Boppart (2014) by introducing a type of ”price independent generalized linearity” (PIGL) preferences. Preferences are specified over the prices P_1 and P_2 and the nominal expenditure level of household j, e_j. The indirect instantaneous utility function is

$$V(P_1, P_2, e_j) = \frac{1}{\rho}\left(\frac{e_j}{P_2}\right)^{\rho} - \frac{\nu}{\theta}\left(\frac{P_1}{P_2}\right)^{\theta} - \frac{1}{\rho} + \frac{\nu}{\theta}, \tag{16}$$

where the parameters govern how the demand for the two goods reacts to changes in relative prices (θ) and the income level (ρ). ν controls the average expenditure shares of goods 1 and 2. We do not make any explicit parameter restrictions at this point.[14] PIGL preferences offer several advantages. First, they generate non-homothetic behavior of households even as incomes become very large. This is relevant given the long time span we consider in our simulation and stands in contrast to, for example, generalized Stone-Geary preferences, which imply that the effect of the income level on consumption patterns tends towards zero for large income values. Second, PIGL preferences allow for easy aggregation. This is particularly relevant in the case of routine and manual workers with their heterogeneous abilities, as we will show below.

[14] Note that with $\theta = \rho = 0$ we obtain the familiar Cobb-Douglas preferences with $V = \log\left(\frac{e_j}{P_2^{\nu} P_1^{1-\nu}}\right)$, while $\nu = 0$ reduces the preferences to a one good economy with CRRA preferences. Boppart (2014) shows that $e_j^{\rho} \geq \frac{1-\rho}{1-\theta} \nu P_1^{\theta} P_2^{\rho-\theta}$ ensures that preferences are valid.

Optimization

All non-ICT and ICT capital is owned by high-skill households. This assumption is motivated by empirical evidence, as we show based on the Survey of Consumer Finances (SCF) in Appendix A.8.[15] Furthermore, we will show in our simulation that changes in capital income are not the driving force of increasing income inequality.

The budget constraint for high-skill households is

$$\underbrace{C_{1,H}P_1 + C_{2,H}P_2}_{e_H} + I_K P_1 + I_{ICT}P_2 = \bar{H}W_H + Kr_K + ICTr_{ICT}, \tag{17}$$

where r_{ICT} is the depreciation-adjusted net interest rate.

Low-skill households are excluded from capital markets and are hand-to-mouth consumers. The budget constraints for routine and manual workers are

$$\underbrace{C_{1,L}P_1 + C_{2,L}P_2}_{e_R} = \bar{R}W_R, \qquad \underbrace{C_{1,M}P_1 + C_{2,M}P_2}_{e_M} = \bar{M}W_M. \tag{18}$$

All types solve an intratemporal optimization problem, choosing between consumption of goods 1 and 2. We apply Roy's identity to the indirect utility function and obtain the Marshallian demand functions, which express demand for consumption goods as a function of prices and expenditures:

$$C_{1,o} = \nu \frac{e_o}{P_1}\left(\frac{P_2}{e_o}\right)^{\rho}\left(\frac{P_1}{P_2}\right)^{\theta}\Delta_o, \qquad C_{2,o} = \frac{e_o}{P_2}\left(1 - \nu\left(\frac{P_2}{e_o}\right)^{\rho}\left(\frac{P_1}{P_2}\right)^{\theta}\right)\Delta_o \tag{19}$$

where $o = H, R, M$. $\Delta_o = \int_0^o \left[e_j/e_o\right]^{\rho} dj$ is a measure of inequality of consumption expenditures across households j within occupation o.

The elasticity of substitution between goods 1 and 2, σ_j, is household-specific and given by[16]

$$\sigma_j = 1 - \theta - \frac{\nu\left(\frac{P_1}{P_2}\right)^{\theta}}{\left(\frac{e_j}{P_2}\right)^{\rho} - \nu\left(\frac{P_1}{P_2}\right)^{\theta}}(\theta - \rho).$$

Only the high-skill households solve an intertemporal optimization problem, choosing expenditures and investment. Denoting by $\beta \in [0,1]$ the discount factor, the value function can be expressed as

$$W(K, ICT) = \max V(P_1, P_2, e_H) + \beta \mathbb{E}\left[W(K', ICT')\right], \tag{20}$$

which we solve subject to the budget constraint eq. (17). Note that in the presence of population growth, this budget constraint enters the optimization problem in per-capita terms, i.e. dividing

[15]Jaimovich et al. (2021) make the same assumption about capital ownership.
[16]See Boppart (2014) for the proof.

by $\bar{H}$. Optimization results in the Euler equations

$$\left(\frac{P_2'}{P_2}\right)^{\rho}\left(\frac{e_H'}{e_H}\right)^{1-\rho}\left(\frac{\bar{H}'}{\bar{H}}\right)^{\rho}=\beta\frac{(P_2'(1-\delta_{ICT}')\mu'+r_{ICT}')}{\mu P_2},$$
$$\left(\frac{P_2'}{P_2}\right)^{\rho}\left(\frac{e_H'}{e_H}\right)^{1-\rho}\left(\frac{\bar{H}'}{\bar{H}}\right)^{\rho}=\beta\frac{(P_1'(1-\delta_K')+r_K')}{P_1} \quad (21)$$

3.4 Market Clearing

The two capital markets clear

$$K=K_1+K_2, \qquad ICT=ICT_1+ICT_2. \quad (22)$$

The markets for consumption goods clear

$$C_1=C_{1,H}+C_{1,R}+C_{1,M}, \qquad C_2=C_{2,H}+C_{2,R}+C_{2,M}. \quad (23)$$

and finally the labor markets clear

$$\bar{H}=H_1+H_2, \qquad \bar{R}=R_1+R_2, \qquad \bar{M}=M_1+M_2 \quad (24)$$

3.5 Equilibrium

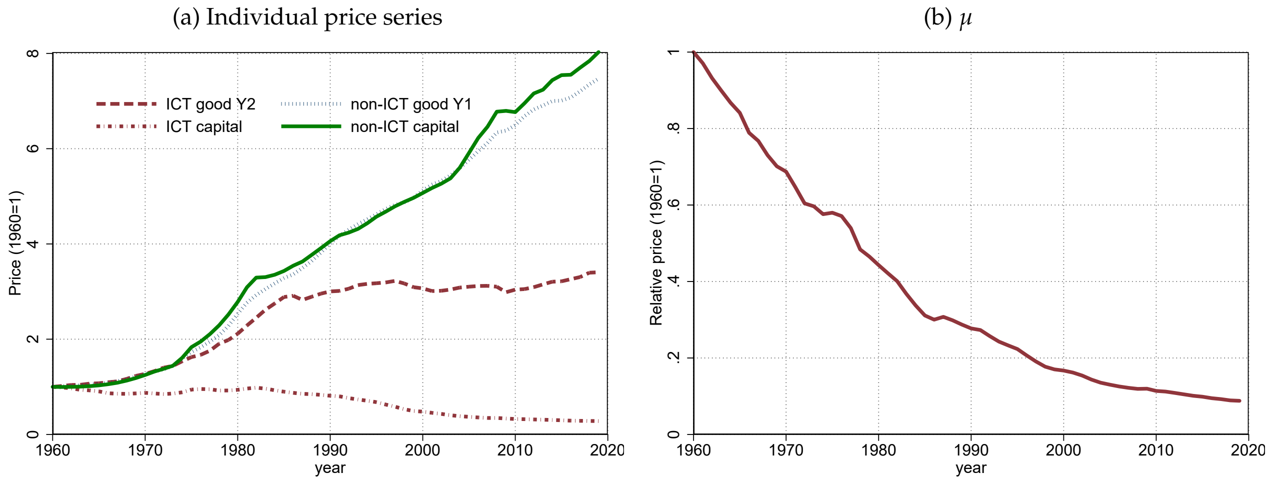
Given a path for μ, the distribution of low-skill abilities $\Gamma(\lambda_R,\lambda_M)$ and the exogenous supply of high- and low-skill workers $\bar{H}$, $\bar{L}$, the equilibrium is defined as a set of wages W_H, W_R, W_M, prices P_1, P_2 and returns r_{ICT}, r_K such that the first order conditions of the firms, eqs. (6), (7), (8), (9), (10), are satisfied, the first order conditions of the households, eqs. (19) and (21) are satisfied and markets clear, eqs. (22), (23) and (24).

4 Calibration and Simulation

We first calibrate all parameters that can be identified directly from the data and set some other parameters in line with the literature. All remaining parameters are internally calibrated using the Simulated Method of Moments (McFadden, 1989). Throughout, we assume that digitalization (jointly with an increase in skill supply) fully explains the changes observed in the data.

In the simulation, we assume that the economy is at a pre-digitalization steady state in 1960. Between 1960 and 2017, digitalization takes place via changes in μ. The model economy converges to a new steady state at the 2017 value of μ afterwards. Along the transition, we assume a perfect foresight equilibrium. In this setting, we assess how digitalization affects consumption of each occupation, and how strong the price and income channels are, respectively.

Figure 9: Prices of final consumption goods and capital



Note: Panel (a) shows Törnqvist price indices based on industry- and asset-level data. Industries are classified into ICT- and non-ICT sector according to the clustering. Assets are classified according to the definition used in Section 2.1. μ in panel (b) is the price series for ICT capital (red dashed-dotted line of panel (a)) relative to the price series for good 2 (red dashed line). *Source:* NIPA and own calculations.

4.1 Externally Chosen Parameters

Advances in Digital Technology

In order to define ICT-intensive industries, we consider our main industry-level ICT intensity measure and use kmeans-clustering to sort the 61 private BEA industries into two groups. For the average years of 1996-1998 ten industries[17] are classified to be part of the ICT-intensive sector that corresponds to our sector 2. Here and in other parts of the calibration, we choose the years around 1996 because this is the first year covered by our consumption dataset, and it lies roughly in the middle of our simulation period. Sector 2 has an ICT intensity of 45% and combines a value added share of 17.3%. Sector 1 has an ICT intensity of 8.3%.[18]

According to eq. (11), progress in ICT technology works through a decline in μ, the price of ICT capital relative to output of good 2. Panel (a) of Figure 9 shows the evolution of the prices of Y_1 and Y_2 as well as the prices of ICT capital (I_{ICT}) and non-ICT capital (I_K). The series are Törnqvist price indices based on NIPA data described in further detail in Appendix A.1. While the prices of Y_1 and I_K move in parallel, justifying the assumption of a constant rate of transformation, there is a strong decline in the price of I_{ICT} relative to Y_2. The resulting μ is shown in panel (b). We feed a smoothed version of this series into our simulation as an exogenous change in ICT technology.

[17] These are: BEA codes 3340 (Computer and electronic products), 5110 Publishing industries, , 5140 Data processing, internet publishing, and other information services , 5230 Securities, commodity contracts, investments, and related activities , 5240 Insurance carriers and related activities , 5250 Funds, trusts, and other financial vehicles , 5411 Legal services , 5412 Accounting and bookkeeping services , 5415 Computer systems design and related services , 5500 Management of companies and enterprises. If we cluster between 2015 and 2017, one more industry is classified as ICT-intensive: 5610 Administrative and support services. As this industry is not very large, the calibration for the two sectors in the model does not change a lot.

[18] Alternatively, we could classify industries according to whether they produce ICT capital. In the capital flow data published by the BEA, the assets that we classify as ICT assets are produced by three industries: 3340, 5110 and 5415. This is a subset of our ten ICT-using industries.

Skill Supply and Occupational Choice

We feed in the supply of high-skill and low-skill labor, $\bar{H}$ and $\bar{L}$ from the Decennial Census and the American Community Survey (see Appendix A.1 for details) as exogenous processes. These change over time, reflecting an increase in the relative supply of high-skill labor during the 60 years of our sample. Low-skill workers decide whether to work in manual or routine occupations based on their individual abilities and the equilibrium wage (eq. 13.) We specify the distribution of routine and manual abilities, $\Gamma(\lambda_R, \lambda_M)$ to be jointly log-normal. There are three parameters to specify: two standard deviations, two means and one correlation. Following Jaimovich et al. (2021), we normalize the means to 1 and assume that the correlation between the two ability draws is zero. We obtain the standard deviations of wages directly from the data. We require the ratio between routine and manual wages W_R/W_M to be such that the model exactly produces the routine worker share observed in the data. Concretely, we feed the supply of routine and non-routine manual work from the data into the model and then use the resulting wage ratio as a data moment we want to target (see below). We also calibrate the income distribution parameters Δ_o that are used for the Marshallian demands. To do so, we create a quantile function (inverse CDF) from the CDF $\Gamma'(\lambda_R, \lambda_M)$. This function then represents the distribution of wages across the skill distribution.

Production Function

The depreciation rates δ_K, δ_{ICT} and the capital shares α_i are taken from the BEA. The non-ICT capital share in those industries can be computed as one minus the labor share, adjusted for the share of non-ICT capital in overall capital in both sectors. The share of manual work ψ_i is also taken directly from the data, multiplying the share of non-routine manual workers from the ACS with the labor share from Elsby et al. (2013).

4.2 Internally Calibrated Parameters

The remaining parameters of the production function and the preference parameters are calibrated via the Simulated Method of Moments (McFadden, 1989). Specifically, these are the four elasticities of substitution (ϵ_i, η_i $\forall i = 1, 2$) and the four weights (γ_i, ϕ_i) in the production function, and ν, ρ and θ in indirect utility function. We make the following simplification: $\epsilon_1 = \epsilon_2 \equiv \epsilon$, $\eta_1 = \eta_2 \equiv \eta$. In the data, the two sectors are at any time characterized by large differences in the proportions to which they use the different input factors, which are captured by differences in ϕ_i and γ_i. Differences in the evolution of input factor use over time are less pronounced, motivating us to choose the same elasticities ϵ and η across sectors.[19]

This leaves us with nine parameters to be calibrated. We aim to match nine moments: The wage premium of high-skill relative to routine workers, the employment shares of routine and manual workers (via the wage premium of routine relative to manual workers), the price of good 1 relative to good 2, the ICT intensity, the differences between high-skill and manual workers in the

[19]See Klump and de La Grandville (2000) for a more general discussion on how weights and elasticities are related.

Table 2: Calibration fit

Data moment	Source	Data	Model
Skill premium, 1960-2017	ACS	37.7	37.5
Relative price, 1960-2017	NIPA	134.6	128.3
ICT intensity, 1960-2017	BEA	2096.2	1893.9
Relative consumption, 1960-2017	NIPA	-79.7	-72.1
Labor share, 1960-2017	EHS (2013)	-11.7	-17.8
Rate of return to ICT capital, 1960-2017	KN (2019)	-86.0	-89.8
Median increase in exp. share good 2, 1996-2017	see Sec. 2	104.7	32.1
Percent diff., exp. share good 2 Top10 Bot 10, 2017	see Sec. 2	40.8	50
Employment share decline R, 1960-2017	ACS	-35.5	-36.4

Note: All moments are expressed in percentage changes, except for the difference in expenditure shares. Note that we match employment shares of all three occupations over time, see Fig. B.1. KN refers to Karabarbounis and Neiman (2019). Given the volatility in the return series, we apply HP filtering. EHS refers to Elsby et al. (2013). We apply HP filtering and extrapolate the series between 2013-2017. The data on the expenditure share is available only from 1996 and taken from Section 2.

expenditure share of good 2, the quantities consumed of good 1 relative to good 2, the labor share, the rate of return on ICT capital and the expenditure share of good 2 for the median household. While the expenditure shares, relative quantities and prices are more relevant for the parameters in the indirect utility function, the remaining moments are primarily used to pin down the parameters in the production function. Most moments are expressed in percentage changes between 1960 and 2017, or between the earliest available year and 2017. The system is exactly identified under the assumption of moments orthogonality. Table 2 compares the targeted data moments with the model. Appendix B.1 provides further details on the calibration procedure, the solution algorithm and how the moments link to the parameters.

Table 3 summarizes the calibrated parameters. The large elasticity of substitution in the inner nest, η, suggests that routine labor and ICT capital are strong substitutes, whereas the composite of these two and high-skilled labor are complementary to each other due to the lower value of ϵ. This is in line with the existing literature (e.g. Eden and Gaggl, 2018, Vom Lehn, 2020 and Jaimovich et al., 2021). We see not surprisingly that $\phi_1 > \phi_2$, as sector 2 is more ICT-intensive whereas sector 1 relies more on routine labor input. In the middle nest, the input generated using these two production factors is more important in sector 2, whereas sector 1 places a high weight on high-skill labor. As sector 1 relies more on non-ICT capital ($\alpha_1 > \alpha_2$) and on manual work ($\psi_1 > \psi_2$), high-skill labor plays a less important role than the large value of γ_1 suggests.

A positive value of ρ indicates that the ICT good is a luxury good, in line with the empirical evidence. $\theta < 0$ implies that when the relative price of one good increases, the share of expenditure spent on this good decreases. This means that the two goods are substitutes. Indeed, the elasticity of substitution between ICT and non-ICT goods for the average household is $\sigma = 1.04$. With an elasticity of substitution above one, the sector that experiences a relative price increase (i.e. the non-ICT sector) shrinks in terms of expenditure shares.[20] To our knowledge, our paper is the first to provide estimates of the elasticity of substitution between ICT and non-ICT goods.

[20]Compare this to Boppart (2014) who considers a goods-services dichotomy and finds these two to be complements.

Table 3: Calibrated parameters

Symbol	Value	Description	Source
Inner nest			
ϕ_1	0.896	Weight of R in 1	calibrated
ϕ_2	0.687	Weight of R in 2	calibrated
$\eta_1 = \eta_2$	2.170	El. of Subst between R and ICT	calibrated
Middle nest 1			
γ_1	0.548	Weight of H in sector 1	calibrated
γ_2	0.217	Weight of H in sector 2	calibrated
$\epsilon_1 = \epsilon_2$	0.651	El. of Subst betw H and AW	calibrated
Middle nest 2			
ψ_1	0.143	Income share of M in sector 1	ACS and EHS
ψ_2	0.021	Income share of M in sector 2	ACS and EHS
Final Production			
α_1	0.45	Capital share in sector 1	BEA
α_2	0.18	Capital share in sector 2	BEA
δ_{ICT}	0.14	Depreciation rate ICT capital	BEA
δ_K	0.06	Depreciation rate non-ICT capital	BEA
Preferences			
ν	0.580	Expenditure share parameter good 1	calibrated
θ	-0.042	Substitution parameter	calibrated
ρ	0.117	Income elasticity parameter	calibrated
β	0.97	Discount factor	

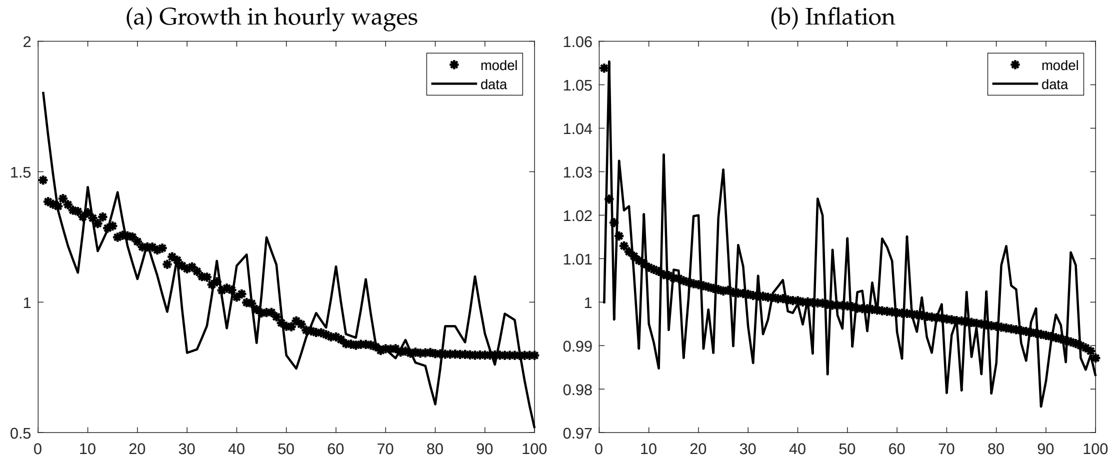
4.3 Validation

We study the validity of the estimated model by contrasting non-targeted model moments with data moments, choosing those that are informative for the fit of the model in important dimensions.

We start by considering the evolution of income across the distribution over the sample period. We simulate 100,000 workers in 1960 by drawing manual and routine ability pairs λ_R, λ_M for each and sorting them into the occupation that pays the highest wage. We then rank them according to their income. We compare this to an income distribution based on 2017 wages, where we keep the ability distribution constant (i.e. following the same individuals over time).

Figure 10 (a) plots the percentage change in (nominal) hourly wages between 1960 and 2017 in the ACS and in the model for low-skill workers. Since price changes differ in the data and in the model – as we have normalized $P_2 = 1$ – we normalize both series to 1 for the mean across all percentiles. The fit between model and data for this untargeted moment is very good. The model captures both the negative slope and the flattening out of the curve towards the upper end of the distribution. To cover the full distribution (also including the high-skill households), we

Figure 10: Wages and inflation across the income distribution



Note: The figure compares model and data for key untargeted moments along the income distribution. Panel (a) shows the growth rate in hourly wages between 1960 and 2017. Panel (b) shows the growth rate in prices between 1996 and 2017. Both series are normalized to 1 for the mean across all percentiles. *Sources: ACS, BLS, CEX and own calculations.*

Table 4: Calibration fit, non-targeted moments

Data moment	Source	Data	Model
value-added	BEA	548.51	794.12
share of value-added falling on sector 1	BEA	-15.34	-27.07
share of ICT capital used in sector 1	BEA	47.82	104.33
share of K capital used in sector 1	BEA	-2.50	-14.60
share of M workers used in sector 1	ACS	-0.39	-8.93
share of H workers used in sector 1	ACS	-12.94	-37.05
share of R workers used in sector 1	ACS	-2.5	24.64

Note: All moments are percentage changes between 1960 and 2017.

compare the change in the growth rate of wages at the top 10% to those at the bottom 10% of the income distribution. This number is 0.824 in the data and 1.078 in the model. When considering the growth rate of disposable per-capita income of high-skill workers instead of just their wages, the number is 0.753, giving us a fairly good match between data and model. Thus, we see a clear U-shape of income changes.

In Figure 10 (b) we consider inflation rates across the income distribution. We plot the data series of Figure 5 (b) against the model equivalent. Again, due to the normalization of prices in the model, the focus is on the relative differences between the income percentiles rather than on the level. The fit between model and data is almost perfect. So the model captures the differences in inflation rates for high- vs. low-income households very well, a prerequisite for getting a good model prediction for the price channel.

Table 4 shows how the model matches additional non-targeted moments mostly relating to the sectoral composition. We match the growth of the aggregate economy fairly well. The share of sector 1 in value-added shrinks in both model and data, although the shift of economic activity is stronger in the model. Similarly, for the use of ICT and non-ICT capital we get the right direction of changes, but larger magnitudes. This is because there are no frictions for sectoral reallocation in the model, whereas in reality the use of factor inputs across sectors does not adjust as easily. With regards to the allocation of workers across sectors, we match reasonably well the movement of manual and high-skill workers over time, as both types shift towards sector 2. For routine workers, an increase in the share working in sector 1 in the model contrasts with a decrease in the data. This may be because adjustment of production processes takes time in the real world, whereas it can happen instantaneously in the model.

5 Results

In this section, we discuss the simulation results in detail, focusing on the effect of digitalization on prices, incomes and consumption of households. We show that price changes vary substantially across the income distribution by comparing inflation rates across households. To estimate the importance of the price effect for welfare, we carry out a counterfactual analysis in form of a compensatory variation, similar to the exercise in the empirical part. To arrive at the welfare effects of digitalization, a U-shaped pattern for income polarization is combined with inflation costs that are strictly downward-sloping in income.

5.1 Price Effect

Progress in digitalization favors sector 2, which uses ICT capital more intensely. As sector 2 produces relatively cheaper, the relative price of good 1 increases. The normalization of the price of good 2 implies consumer-price inflation over the considered time horizon[21]. At the same time, heterogeneous workers in our model are characterized by different consumption bundles which means they face different individual inflation rates. We study changes in the costs of living for each of the three occupations and then for the different individuals in the group of low-skilled workers (both routine and manual). We calculate Törnqvist price indices, on the basis of which we derive occupation-specific and individual inflation rates.

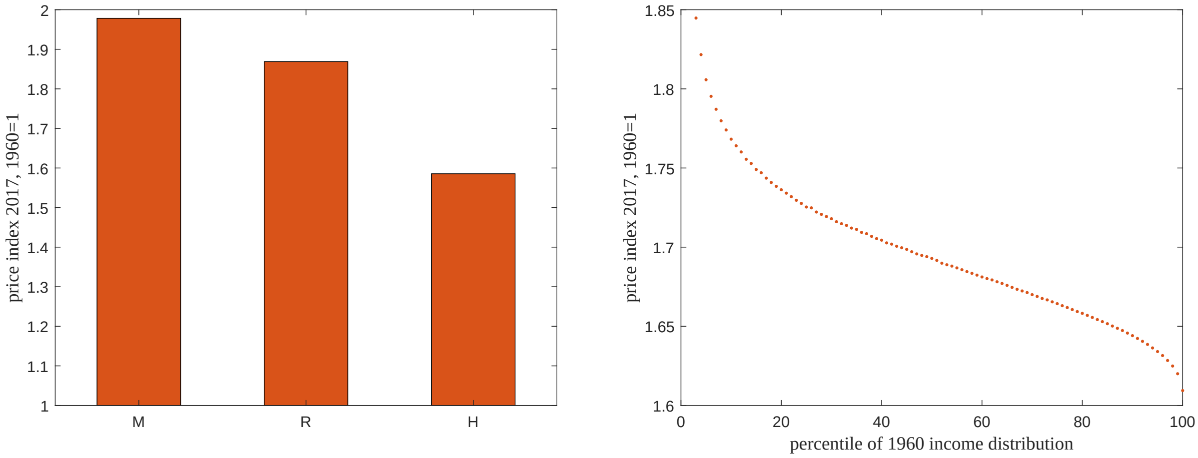
Figure 11 (a) shows the occupation-specific consumer price index in 2017 compared to 1960. The costs of living increase for all groups, but importantly, the price index for high-skill households is lower than for the other two groups as their consumption bundle depends less on non-ICT goods. The differences between manual and routine workers are considerably smaller. Figure 11 (b) studies this group in more detail. As each income percentile has slightly different consumption baskets, their inflation rates differ as well: The price index for the richest in the group is almost 20 percentage points lower than for the poorest.

[21] We do not aim to match the aggregate inflation rate which is a monetary phenomenon. We rather focus on differences in inflation between households that arise due to changes in relative prices caused by technological progress.

Figure 11: Price indices

(a) by group

(b) by percentile



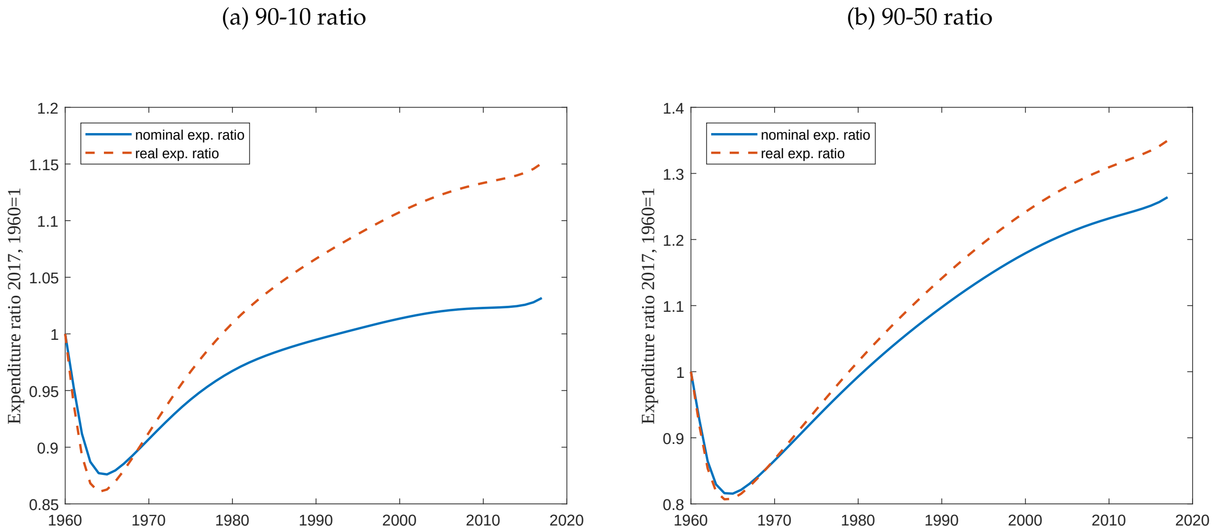
Note: Panel (a) shows the Törnqvist price indices in 2017 (1960=1) for the manual, routine and high-skill worker groups. Panel (b) shows the Price index in 2017 for percentiles of the income distribution of low-skill (manual and routine) workers. Inflation rates for percentile one and two are outliers and lie above 2.

5.2 Consumption Inequality

We now focus on changes in consumption inequality. Figure 12 presents two commonly used measures of inequality, the ratio of the 90th relative to the 10th percentile (90-10 ratio) and the ratio of the 90th relative to the 50th percentile (90-50 ratio) of the income distribution. The solid blue line is based on nominal expenditures, which is the measure generally considered in the consumption inequality literature. Both ratios increase over time, but the increase in the 90-10 ratio is much smaller. Our numbers roughly align with Meyer and Sullivan (2023), who document an increase in the 90-10 ratio by 9.5% over the same time horizon, and an increase in the 90-50 ratio by 12.85%. Other papers such as Aguiar and Bils (2015) and Attanasio and Pistaferri (2014) report larger numbers, but agree on the general finding that there has been an increase in (nominal) consumption inequality.

When deflating expenditures with the percentile-specific Törnqvist index (dashed red line), Consumption inequality increases by a larger margin in both panels. This is driven by higher price increases for the poor as documented in the previous section. The difference is particularly large for the 90-10 ratio, as inflation rates differ more strongly between the top and bottom of the income distribution. The 90-10 ratio increases by 15% when taking relative price changes into account, compared to just 3% when considering nominal expenditures. This shows that it is important to consider real units when discussing consumption inequality. To households, it is not nominal expenditures that matter, but what they can actually buy with the money they spend. As inflation rates differ across the income distribution, so does purchasing power.

Figure 12: Expenditure ratio over time



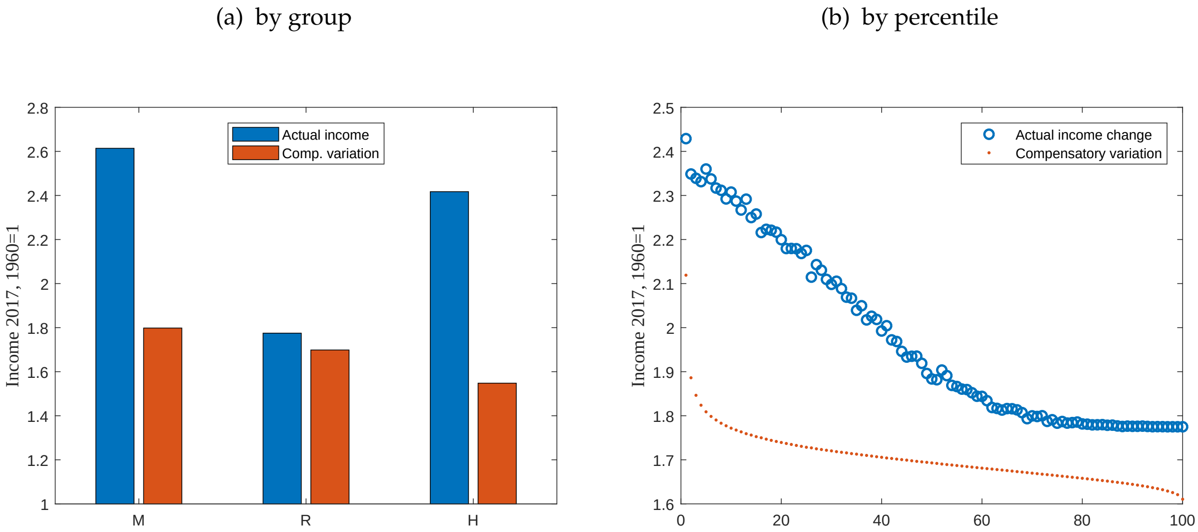
Note: Panel (a) shows the expenditure ratio over time between the 90th and 10th income percentile. Panel (b) shows the expenditure ratio over time between the 90th and 50th income percentile. Nominal expenditure ratios (solid blue line) between percentile 90 and j are defined as exp_{90t}/exp_{jt} and normalized to one in 1960. Real expenditures (dashed red line) are deflated by the percentile-specific price index. The initial dip at the beginning of the transition period is due to temporary reallocation of expenditures from consumption towards investment, as households foreshadow the advances in ICT technology.

5.3 Welfare Effects of Digitalization

We now assess the welfare effects of digitalization via a counterfactual exercise. We conduct a compensatory variation, which asks how much additional income households need to be given in 1960 to be compensated for the increase in the relative price between 1960 and 2017. We then compare this to the actual income increase over the sample period. If the actual income increase is larger than the compensation, households are better off in the post-digitalization world. We thus compute how much different individuals benefit from digitalization and compare the effects across groups. This procedure is analogous to our empirical analysis in Section 2.3, but avoids the shortcomings of the pure data analysis.

Figure 13 (a) compares the compensatory income with the actual income for manual, routine and high-skill workers. The initial income is normalized to 1 for each type. This means that the numbers in the graph are to be interpreted relative to 1960 income. The red bar shows the compensatory income. Both manual and routine workers rely more on consumption of good 1 than high-skill workers, and therefore require a larger compensatory income (manual workers: 180%, routine workers: 170%, compared to 155% for high-skill workers). This means that the price changes are consequential for welfare across the three groups. The blue bar shows how much more income workers actually receive in 2017 relative to 1960. Digitalization has the largest positive effect on the wage of manual workers, whose income increases to 261% of the initial value. This number is 242% for high-skill workers and 178% for routine workers. High-skill workers also experience an increase in their capital income, this increase is however not consequential in this comparison. When considering capital income for high-skill workers as well, the overall income increase goes from 242 % for wages only to 244%.

Figure 13: Income changes and compensatory variation



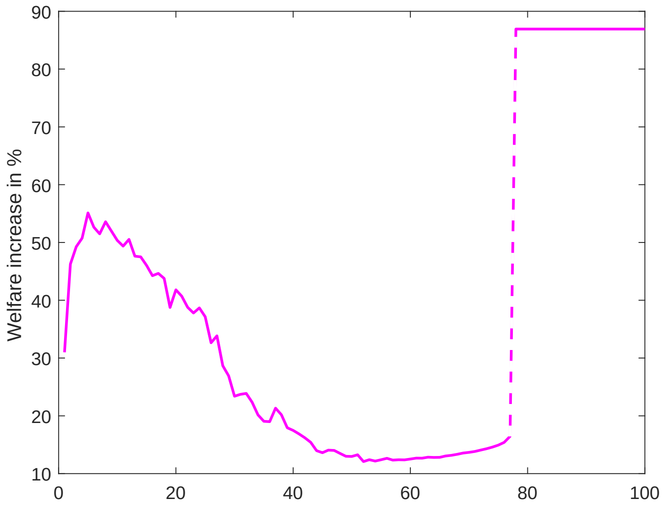
Note: Panel (a) shows actual income in 2017 (blue) and compensatory variation (red) for the manual, routine and high-skill worker groups. Panel (b) shows actual income increase and the compensatory variation for percentiles of the income distribution of low-skill (manual and routine) workers. Income values are relative to 1960 which is normalized to one.

Next, we study the group of low-skill workers in more detail. In contrast to high-skill workers, who are homogeneous in the model, there are substantial differences in abilities, and thus in income, across low-skill workers, both across, but also within manual and routine occupations. The actual income distribution plotted Figure 13 (b) is the same as shown in Figure 10. The compensatory variation is strictly decreasing along the income distribution, reflecting the increasing share of ICT-intensive consumption and the large differences in individual inflation rates across households. For the middle of the income distribution, especially the percentiles 60-80 of low-skill households, much of the income increase is eaten up by rising prices.

This gets particularly clear when considering the net welfare effects, deducting compensatory income from actual income. As illustrated in Figure 14, the overall gains from digitalization are very unequally distributed. High-skill households –which constitute the top 23% in the initial 1960 distribution– gain welfare equivalent to 87% of their initial income, and are therefore the main beneficiaries from digitalization. The original 242 % income increase reduces to 87% when including price changes. The poorest decile, which experiences a substantial income gain, still benefits, however some of their income increases get eaten up by price increases of their consumption basket. A 235 % increase in income translates into a 50% increase in welfare only. Even though both experienced a similar increase in income, their welfare increase diverged substantially because of the price effect. The worst off group are the percentiles 40-77, which hardly gain from digitalization at all, with gains slightly above 10 %. Overall, the U-shaped income effect is set against strictly downward sloping inflation costs. Taken together, these two channels suggest a welfare effect that looks rather J-shaped than U-shaped.[22] While the price channels counteracts the relatively higher income for the poor, the overall welfare effect is amplified for the top as price changes work more in their favor.

[22]If the group of high-skill workers were heterogeneous, the effect would be slightly more nuanced, although the overall pattern would still prevail.

Figure 14: Welfare effects: Actual income increase - compensatory variation



Note: The figure shows the income equivalent welfare increase in 2017 compared to 1960 in %. The welfare increase is computed by subtracting the compensatory variation from the actual income increase. The size of low skill groups is compressed such that 23 % of the population count as High-skill, in line with the initial distribution in 1960. The dashed line connects the two groups.

6 Conclusion

This paper discusses new ways to measure the impact of digitalization on consumption, prices and time use. Our results point out that the effect of digitalization on consumption inequality is even stronger than the substantial increase in income inequality suggests. In providing a first estimate of the size of the price channel, this paper offers a starting point for more follow-up research on the effect of digitalization on welfare and inequality: In how far does digitalization affect quality differences between relatively similar products consumed by poor or the rich? Is there a systematic difference in time allocation for digital products that are for free compared to the ones that are not? In how far does digitalization help to reduces waiting time and the facilitation of transactions for certain income groups? Finally, our paper compares current manual (routine) workers with 1960 manual (routine) workers, i.e. we do not take into account the role of occupation switching.

References

Acemoglu, D. and P. Restrepo (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review 108*(6), 1488–1542.

Adão, R., M. Beraja, and N. Pandalai-Nayar (2020). Technological transitions with skill heterogeneity across generations. *NBER Working Paper No. 26625*.

Aghion, P., C. Antonin, X. Jaravel, and S. Bunel (2020). What are the labor and product market effects of automation? New evidence from France. *CEPR Discussion Paper 14443*.

Aguiar, M. and M. Bils (2015). Has consumption inequality mirrored income inequality? *American Economic Review 105*(9), 2725–2756.

Aguiar, M. and E. Hurst (2005). Consumption versus expenditure. *Journal of Political Economy 113*(5), 919–948.

Akerman, A., I. Gaarder, and M. Mogstad (2015). The skill complementarity of broadband internet. *Quarterly Journal of Economics 130*(4), 1781–1824.

Argente, D. and M. Lee (2021). Cost of living inequality during the great recession. *Journal of the European Economic Association 19*(2), 913–952.

Attanasio, O. P. and L. Pistaferri (2014). Consumption inequality over the last half century: some evidence using the new PSID consumption measure. *American Economic Review 104*(5), 122–26.

Attanasio, O. P. and L. Pistaferri (2016). Consumption inequality. *Journal of Economic Perspectives 30*(2), 3–28.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review 103*(5), 1553–1597.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in US wage inequality: Revising the revisionists. *Review of Economics and Statistics 90*(2), 300–323.

BEA (2021). Updated digital economy estimates - June 2021. *BEA Report*.

Boppart, T. (2014). Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences. *Econometrica 82*(6), 2167–2196.

Borusyak, K. and X. Jaravel (2021). The Distributional Effects of Trade: Theory and Evidence from the United States. Technical report, Available at SSRN, no. 3269579.

Brynjolfsson, E., A. Collis, W. E. Diewert, F. Eggers, and K. J. Fox (2019). GDP-B: accounting for the value of new and free goods in the digital economy. *NBER Working Paper No. 25695*.

Brynjolfsson, E., A. Collis, W. E. Diewert, F. Eggers, and K. J. Fox (2020). Measuring the impact of free goods on real household consumption. *AEA Papers and Proceedings 110*, 25–30.

Burstein, A., E. Morales, and J. Vogel (2019). Changes in between-group inequality: computers, occupations, and international trade. *American Economic Journal: Macroeconomics 11*(2), 348–400.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2017). Disappearing routine jobs: Who, how, and why? *Journal of Monetary Economics 91*, 69–87.

Coyle, D. and D. Nguyen (2020). Free goods and economic welfare. *ESCoE Discussion Paper 2020-18*.

Eden, M. and P. Gaggl (2018). On the welfare implications of automation. *Review of Economic Dynamics 29*, 15–43.

Elsby, M. W. L., B. Hobijn, and A. Sahin (2013). The decline of the US labor share. *Brookings Papers on Economic Activity 2013*(2), 1–63.

Fang, L., A. Hannusch, and P. Silos (2021). Luxuries, Necessities, and the Allocation of Time. *Working Paper*.

Fisher, J. D., D. S. Johnson, J. Marchand, T. M. Smeeding, and B. B. Torrey (2008). The retirement consumption conundrum: Evidence from a consumption survey. *Economics Letters 99*(3), 482–485.

Gaggl, P. and G. C. Wright (2017). A short-run view of what computers do: Evidence from a UK tax incentive. *American Economic Journal: Applied Economics 9*(3), 262–94.

Goos, M. and A. Manning (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *Review of Economics and Statistics 89*(1), 118–133.

Goos, M., A. Manning, and A. Salomons (2009). Job polarization in Europe. *American Economic Review 99*(2), 58–63.

Graetz, G. and G. Michaels (2018). Robots at work. *Review of Economics and Statistics 100*(5), 753–768.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997). Long-Run Implications of Investment-Specific Technological Change. *American Economic Review 87*(3), 342–362.

Hémous, D. and M. Olsen (2022). The rise of the machines: Automation, horizontal innovation and income inequality. *American Economic Journal: Macroeconomics 14*(1), 179–223.

Hobijn, B. and D. Lagakos (2005). Inflation inequality in the United States. *Review of Income and Wealth 51*(4), 581–606.

Horrowitz, K. J. and M. A. Planting (2009). Concepts and methods of the U.S. Input-Output Accounts. *BEA* (September 2006), 1–266.

Hubmer, J. (2023). The race between preferences and technology. *Econometrica 91*(1), 227–261.

Jaimovich, N., I. Saporta-Eksten, H. E. Siu, and Y. Yedid-Levi (2021). The macroeconomics of automation: Data, theory, and policy analysis. *Journal of Monetary Economics 122*, 1–16.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the US retail sector. *Quarterly Journal of Economics 134*(2), 715–783.

Kaplan, G. and S. Schulhofer-Wohl (2017). Inflation at the household level. *Journal of Monetary Economics 91*, 19–38.

Karabarbounis, L. and B. Neiman (2014). The global decline of the labor share. *Quarterly Journal of Economics 129*, 61–103.

Karabarbounis, L. and B. Neiman (2019). Accounting for factorless income. *NBER Macroeconomics Annual 33*(1), 167–228.

Klump, R. and O. de La Grandville (2000). Economic growth and the elasticity of substitution: Two theorems and some suggestions. *American Economic Review 90*(1), 282–291.

Koh, D., R. Santaeulalia-Llopis, and Y. Zheng (2020). Labor share decline and intellectual property products capital. *Econometrica 88*(6), 2609–2628.

Krueger, D. and F. Perri (2006). Does income inequality lead to consumption inequality? Evidence and theory. *The Review of Economic Studies 73*(1), 163–193.

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*(5), 1029–1053.

Kuhn, M. and J.-V. Ríos-Rull (2016). 2013 update on the US earnings, income, and wealth distributional facts: A view from macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly Review 37*(1), 2–73.

Mann, K. and L. Püttmann (2023). Benign effects of automation: New evidence from patent texts. *Review of Economics and Statistics 105*(3), 580–595.

Martinez, J. (2021). Putty-clay automation. *CEPR Discussion Paper 16022*.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica 57*(5), 995–1026.

Meyer, B. D. and J. X. Sullivan (2023). Consumption and income inequality in the United States since the 1960s. *Journal of Political Economy 131*(2), 247–284.

Michaels, G., A. Natraj, and J. van Reenen (2014). Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years. *The Review of Economics and Statistics 96*(1), 60–77.

Moll, B., L. Rachel, and P. Restrepo (2022). Uneven growth: Automation's impact on income and wealth inequality. *Econometrica 90*(6), 2645–2683.

Prettner, K. and H. Strulik (2020). Innovation, automation, and inequality: Policy challenges in the race against the machine. *Journal of Monetary Economics 116*, 249–265.

Villaverde, J. F. and D. Krueger (2007). Consumption over the life cycle : Facts from Consumer Expenditure Survey data. *Review of Economics and Statistics 89*(3), 552–565.

Vom Lehn, C. (2020). Labor market polarization, the decline of routine work, and technological change: A quantitative analysis. *Journal of Monetary Economics 110*, 62–80.

A Data Appendix

A.1 Details on Data Sources and Dataset Construction

Input-Output Accounts

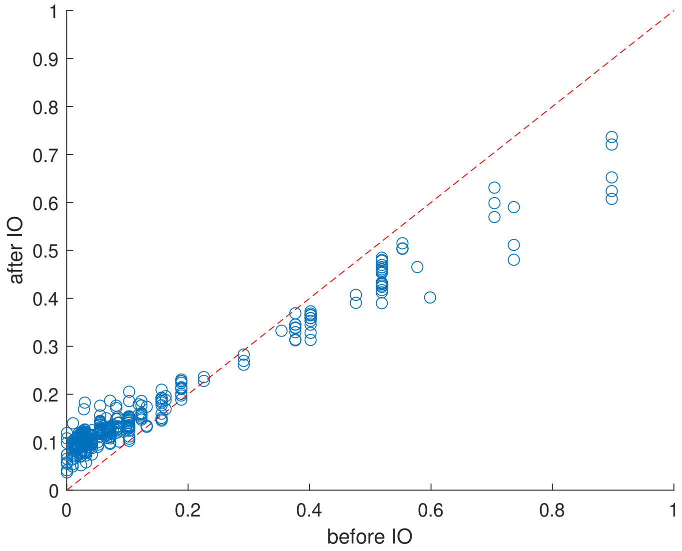
We use the BEA's Input Output Accounts and focus on the detailed Input-Output tables after redefinitions. We use producer-value tables, which means that the distribution margin (i.e. the cost of wholesaling, retailing and transportation) is modeled as a flow from the distribution industries to the final consumer rather than as a flow from the producing industry to the final consumer. This is the standard approach and seems appropriate here as we have no reason to believe that the digitalization content of distribution differs between the different goods and services.

The Input-Output accounts are presented in a set of different tables, among them *use, make* and *direct requirements* tables. We start from the commodity-by-industry direct requirements table, which shows the amount of each commodity that is required by an industry to produce one dollar of the industry's output. The problem with this table is that the same commodity can be produced by different industries, e.g. ice-cream can be produced by the dairy product manufacturing industry and the ice-cream manufacturing industry. We follow Horrowitz and Planting (2009) in creating a commodity-by-commodity direct requirements matrix that takes the market shares of each industry in producing certain commodities into account. With this matrix, we can calculate the digitalization share of final goods and services as a weighted mean of the digitalization shares of its intermediate inputs and of value added.

We pursue an iterative approach. We initially assign to each commodity the digitalization share of the industry that is the ultimate producer of the commodity. Then, we consider all commodities that use this specific commodity as intermediate input. We update the output digitalization share by calculating a weighted mean of the digitalization shares of all inputs and the digitalization share of value added of the final producer. We again assign this share to the inputs used to produce other commodities, and update the output digitalization share again, etc. We continue this procedure until the commodity digitalization shares have converged to fixed values.

To illustrate the importance of the input-output structure, Figure A1 shows the ICT shares for 2012 before and after taking input-output linkages into account. Commodities that initially have a low degree of digitalization when considering only the ICT intensity of the final goods-producing industry tend to be more digitized after the inputs from other industries have been considered. They lie above the dotted red 45 degree line. The reverse is true for very digitized industries that also use inputs from less digitized industries.

Figure A1: ICT intensity with and without considering the IO structure



Note: The graph shows ICT intensities for 2012 before the IO structure has been taken into account (x-axis) and afterwards (y-axis). Each dot represents an IO industry. The red dotted line is the 45-degree line. *Source:* BEA and own calculations.

CEX

We convert all purchases into annual values at constant 2010 US Dollars. We measure income as total before-tax household income, which corresponds to FINCBTAX (FINCBTXM in later vintages) for the interview survey and FINCBEFX (FINCBEFM) for the diary survey. The main component in this measure is labor income. For the group of households we consider – households that participate in the labor market and are not self-employed – labor income makes for around 94% of total income on average (less at lower points of the income distribution because of the higher dependence on social security). Next to labor income, the income measure includes farm and non-farm business income, social security income, interest on savings accounts or bonds, income from dividends, royalties, estates and trusts and rental income. We have carried out robustness checks focusing more narrowly on labor income, which is captured by the variable FSALARYX (FSALARYM in later vintages) in the interview survey and FWAGEX (FWAGEXM) in the diary survey. The results are very similar and are available upon request. Both income and expenditure in the CEX is at the household level. We create individual-equivalent observations by dividing the values by the square root of the number of household members and multiplying the sample weights with the number of household members (see e.g. Villaverde and Krueger, 2007; Fisher et al., 2008). This is important because households at different points in the income distribution vary in their size. In particular, poorer households tend to have more children. The average household size in the bottom quintile of the income distribution is 3.6, whereas it is 2.3 at the top quintile.

We drop households that do not stay in the survey for the entire four quarters of the interview or two weeks of the diary. We consider only households where the head is between 16 and 64 years old and in the labor force, and we drop self-employed. This is because we want to focus on households where labor income is the main source of income. In defining the head, we divert from the CEX convention by making the head the man in mixed couples. We drop the top and bottom 5% of the income distribution in order to mitigate the effect of outliers and top-coding.

The CEX is known for under-reporting of expenditure, in particular by richer households (Aguiar and Bils, 2015; Attanasio and Pistaferri, 2016). This problem has been increasing over time. In consequence, inequality measures like the total consumption expenditure of rich relative to poor households are biased downwards. For the purpose of this analysis, we consider the spending of households of a specific income bin on different products relative to their total expenditure. As long as rich households underreport all expenditures to an equal extent, our results should not be biased.

We observe only total expenditure on each item, and cannot separate between prices paid and quantities purchased. This is problematic, because households may face different prices for the same good, or choose differently-priced products within the narrow categories of the CEX. This issue has been recognized by the literature on consumption inequality (see Attanasio and Pistaferri (2016) for a discussion) and means that we are potentially overestimating consumption inequality.

One dimension of inequality that we cannot capture with our data are quality differences in the goods consumed. A rare exception is the sub-category restaurants (broad industry 722, corresponding to "food away from home" in CEX). Households in the CEX are asked to report separately money spent at fast food, take-out, delivery and vending machines, at full-service restaurants, and at the workplace. Households in the lowest decile spend about 60% of their total restaurant expenses on fast food, but only 30% in full-service restaurants. In contrast, households in the highest decile spend 40% of their total restaurant expenses on fast food, but 55% in full-service restaurants. As full-service restaurants are presumably more labor-intensive than fast food restaurants, the restaurant expenses of high-income households will likely be less automation-intensive than the restaurant expenses of low-income households. Advances in digitalization technology will thus affect them differently.

ATUS

The ATUS survey usually has between 10,000 and 20,000 respondents and differentiates between more than 400 different activity categories. ATUS is addressed to individuals, not households. We use all respondents in the survey and consider their overall household income from the CPS. This makes the measure of time use more comparable to our analysis regarding expenditures, as the CEX reports on the household level. We are interested in the respondents' activities during their non-work time. This includes home production as well as leisure activities. We exclude from our analysis the time when individuals are working and sleeping. We also exclude the use of government services from our analysis as we lack an ICT measure for this sector. ATUS asks survey participants about their primary activities within 15 minute windows. Participants can report more than one primary activity. If they were for example playing games on their smartphone while traveling on the train, they could respond that the primary activity was traveling, but alternatively could also respond that their primary activity was both traveling and playing games. In the latter case, the 15 minutes are equally distributed to both activities.

We develop a concordance that matches ATUS activity codes to IO categories. AsFang et al. (2021), we assume that most activities have a product associated with them, which then allows us to assign an ICT intensity for this activity. For example, we match the activity "laundry" (ATUS-code: 20102) to "household laundry equipment" (IO-code 335224). Some categories are relatively broad in the ATUS, for example "grocery shopping" (ATUS 70101) or "food and drink preperation" (ATUS-code 20201) while the IO level distinguishes between fruits, bread, milk etc. In those cases, we summarize the IO categories to the next higher aggregate IO-level, which would be in this case "food manufacturing" (IO Code 311). In some cases ATUS is more precise, the survey distinguishes for example between travel time to different activities (first two digits in ATUS code: 18) whereas the IO category that we match to is transportation (480000). While Fang et al. (2021) match broad ATUS categories to broad CEX expenditure categories, our concordance is as disaggregate as possible and uses IO commodities as a match.

Census and ACS

Most of our empirical measures of employment and wages are constructed using data from the Decennial Census (for the years 1960, 1970, 1980 and 1990) and the American Community Survey (ACS, available annually from 2000). For the years in between, we interpolate the data. In the sample composition as well as the definition of routine and non-routine workers, we follow Autor and Dorn (2013). We keep only individuals between ages 16-64 that are employed and non-institutionalized. Our key measures are the number of hours worked per year and the hourly wage. The number of hours worked is the number of hours usually worked during a week times the number of weeks worked within a year. As income is top-coded, we replace income in the top bracket with 1.5 times the actual value and then divide by the number of hours worked to obtain the hourly wage. We report all Dollar values in constant 1999 US Dollars. We create a consistent occupational classification using the crosswalks provided on David Dorn's website. We classify the following occupations as routine: sales and administrative support, precision production and craft workers and machine operators. We classify as non-routine cognitive workers managers, professionals and technicians. Non-routine manual occupations are those working in construction and mining, services, mechanics, repairers and transportation. We drop individuals working in the agricultural sector or the military. Out of the total of 318 occupations, 210 are routine or non-routine manual. These are the ones that we focus on when modeling occupational choice. When classifying workers by education, we define high-skill workers as those with at least a college degree and low-skill workers as all others.

NIPA

The goods and assets price series in Figure 9 are constructed using data from NIPA. For the price series, we use NIPA Tables 2.5.3, 2.5.4 and 2.5.5 to construct series for consumed quantities, prices and personal consumption expenditures at the industry level. We then apply our mapping of BEA industries to the ICT- and non-ICT sectors of the model as further described in section 4.1 and create Törnqvist price indices for each sector. For the asset series, we use NIPA Tables 5.4.3, 5.4.4 and 5.4.5, which contain quantities, prices and expenditures for structures, Tables 5.5.3, 5.5.4 and 5.5.5 for equipment, and Tables 5.6.3, 5.6.4 and 5.6.5 for intellectual property products. We merge all three and proceed as with the goods price series to create Törnqvist price indices for the ICT- and non-ICT sector in our model.

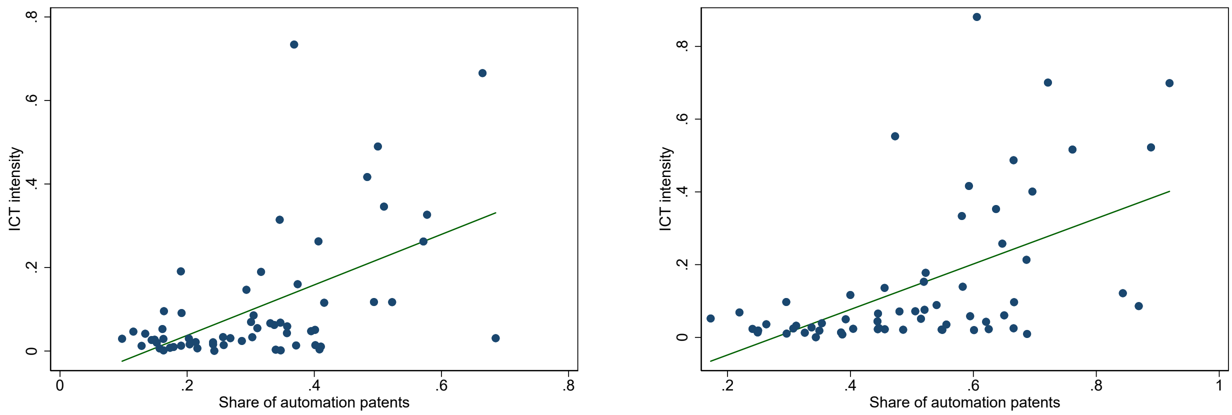
A.2 Comparison of ICT intensity to established measures of digitalization

This section compares our ICT intensity measure with other measures established in the literature. The first comparison is with automation patent data from Mann and Püttmann (2023). Mann and Püttmann (2023) classify US patents as automation or non-automation patents via a text search algorithm and assign patents to the industry of their likely use. While this measure defines automation more broadly, also including robots, computers and communication technology make for the largest share of automation patents. In Figure A2, we plot the correlation of the share of automation patents to our ICT intensity measure for 1990 and 2010. The positive correlation hints that actual investment in ICT capital mirrors new automation technology, which makes us positive that our measure is picking up technological progress in ICT.

Another established procedure in the literature is to consider the input of different tasks across occupations in each industry. Gaggl and Wright (2017) and Adão et al. (2020), among others, argue that in the context of ICT, it is most relevant to focus on cognitive tasks, which are presumably complementary to ICT capital. We define a cognitive-task intensity by taking the log of abstract tasks divided by routine and manual tasks by occupation and defining a job as cognitive-intensive

Figure A2: Relationship between share of automation patents and ICT-intensity measure in 2010

(a) 1990 (b) 2010



Note: The graph plots ICT intensity against the share of automation patents in the total number of patents by BEA industry. Each dot represents one industry, the line shows a linear prediction. The left panel shows raw data for 1990, the right panel for 2010. *Source:* Mann and Püttmann (2023), BEA and own calculations.

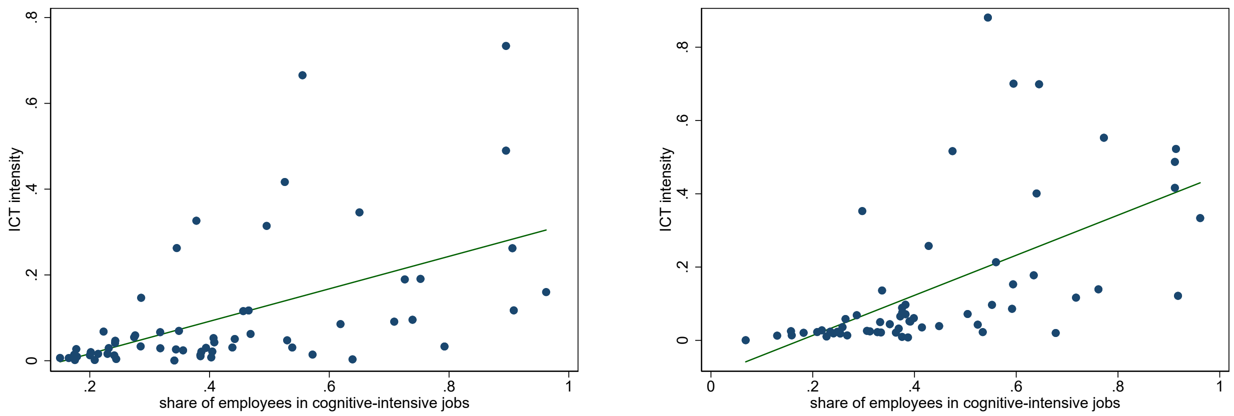
when this number is in the upper third of the distribution.

Figure A4 plots the relationship between the share of cognitive tasks and our ICT intensity measure for 1990 and 2010. There is a clear positive correlation, which increases over time. A higher ICT intensity coincides with a larger share of employees in cognitive-intensive jobs.

Some papers use a similar approach to ours by considering the capital stock at the industry level. Hubmer (2023) defines an equipment intensity as the ratio of equipment and software capital to total capital. Aghion et al. (2020) use a similar approach for France. The measure by Hubmer (2023), although based on the same dataset as ours, differs from ours in several ways: He includes machinery equipment and intellectual property products in the numerator, which are to a large extent classified as non-ICT in our approach. Second, he includes buildings (and the value of land) in the overall capital stock, which we exclude. As a result, the correlation of our measure and Hubmer (2023) on an industry level is almost zero over the considered time horizon. It increases however over time, as ICT becomes a larger share of equipment.

Figure A3: Relationship between share of cognitive employment and the share of ICT capital on overall capital

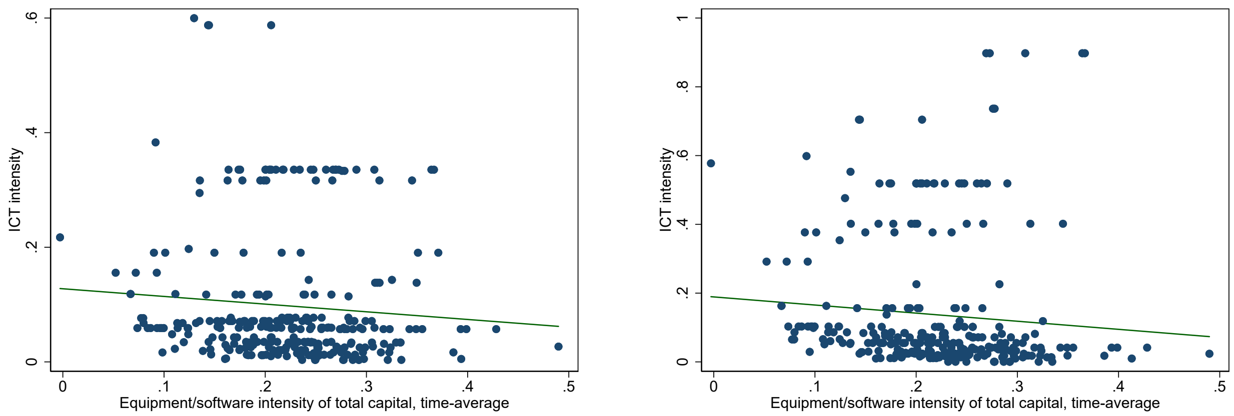
(a) 1990 (b) 2010



Note: The graph plots ICT intensity against the share of employees that work in cognitive-intensive jobs by BEA industry for 1990 (left panel) and 2010 (right panel). Each dot represents one industry, the line shows a linear prediction. *Source:* American Community Survey, Autor and Dorn (2013), BEA and own calculations.

Figure A4: Relationship between the share of overall equipment and software on capital as in Hubmer (2023) and the share of ICT capital on overall capital

(a) 1992 (b) 2012



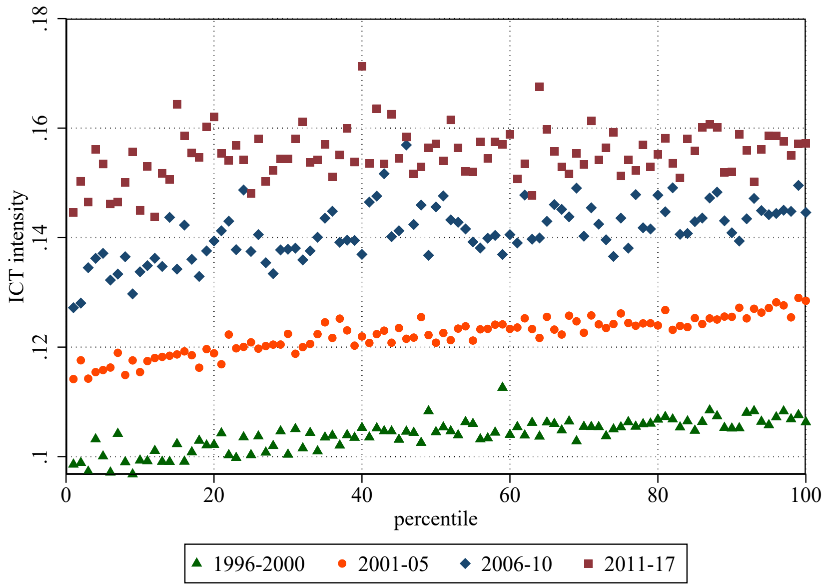
Note: The graph plots ICT intensity against the intensity of equipment and software by BEA industry for 1992 (left panel) and 2012 (right panel). Each dot represents one industry, the line shows a linear prediction. *Source:* Hubmer (2023), BEA and own calculations.

A.3 ICT Intensity in the Consumption Basket: Extensions and Robustness

Non-durable Goods

Consumption in most macroeconomic models can best be understood as consumption of non-durable goods, therefore we exclude in this extension spending on durable goods such as housing, vehicles or education. Figure A5 is the equivalent to Figure 4 but excluding durable goods. This excludes 37 categories from the basket, corresponding to around 30% of expenditures. Among those dropped are expenditures for health and education. We follow Aguiar and Bils (2015) in the non-durable/durable classification. The ICT intensity still increases along the income distribution, but the pattern is slightly weakened.

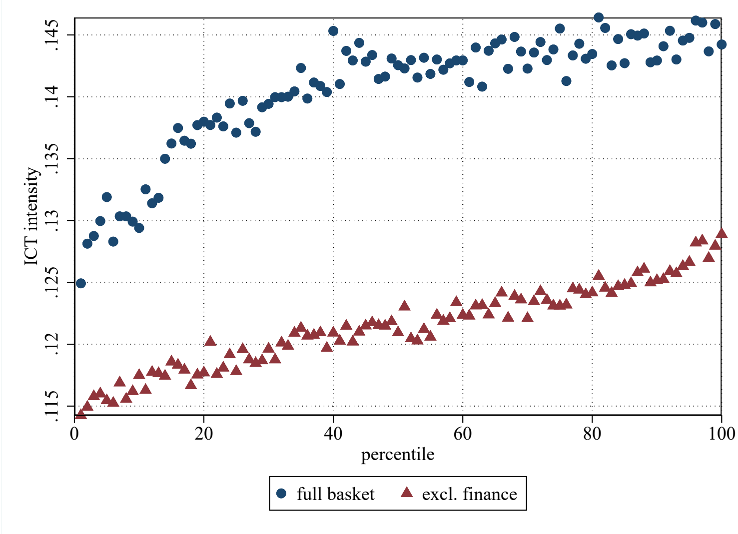
Figure A5: ICT intensity along the income distribution, only non-durable goods



Note: The graph shows the ICT share of the consumption basket by percentile for 2005-2017 period based only on non-durable goods.*Source:* BEA, CEX and own calculations.

Excluding Finance and Insurance

Figure A6: ICT intensity along the income distribution, excluding finance and insurance



Note: The graph shows the ICT share of the consumption basket by percentile for the full sample period including and excluding finance and insurance. *Source:* BEA, CEX and own calculations.

Controlling for Household Age

Household income tends to increase with age, which is why we might be concerned that rather than capturing an increasing ICT intensity of consumption along the income distribution, we capture higher spending shares on digital goods among the age distribution. To refute this concern, we run regressions of the ICT intensity on income, controlling for the age of the household head. In Table A1, the first two columns represent regression results at the level of income percentiles, where income and age are averages across all households within the income percentile. The last three columns show results for regressions at the individual household level. In column (5), we add additional household characteristics, which include the share of children among household members, the sex of the household head, the education of head and spouse and whether each of them participates in the labor market, whether there are earners in the household other than head and spouse, and finally whether the household is located in an urban or rural area. As we cannot match households across the diary and interview survey, we restrict our sample to the interview survey in columns (3) – (5).

In both samples, the coefficient on income is highly significant and remains significant once age is included. As expected, the magnitude becomes is slightly smaller.

Table A1: Regression of ICT intensity on income and age

	Percentiles		Households		
	(1)	(2)	(3)	(4)	(5)
income	0.015***	0.009***	0.021***	0.018***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
age		0.058***		0.032***	0.036***
		(0.006)		(0.001)	(0.001)
Constant	8.378***	6.284***	9.980***	8.822***	9.409***
	(0.071)	(0.241)	(0.058)	(0.072)	(0.103)
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	Yes
Observations	2,200	2,200	70,420	70,420	70,420
R-squared	0.916	0.919	0.389	0.396	0.410

Note: Income is in thousands of constant 2015 USD. ICT intensity is in %. Household-level controls include the share of children, sex of the head, education and labor market participation of head and spouse, presence of other earners, geographical location. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. (3), (4) and (5) are weighted regressions using the survey weights.

A.4 Input- vs. Output-based Digitalization Measures

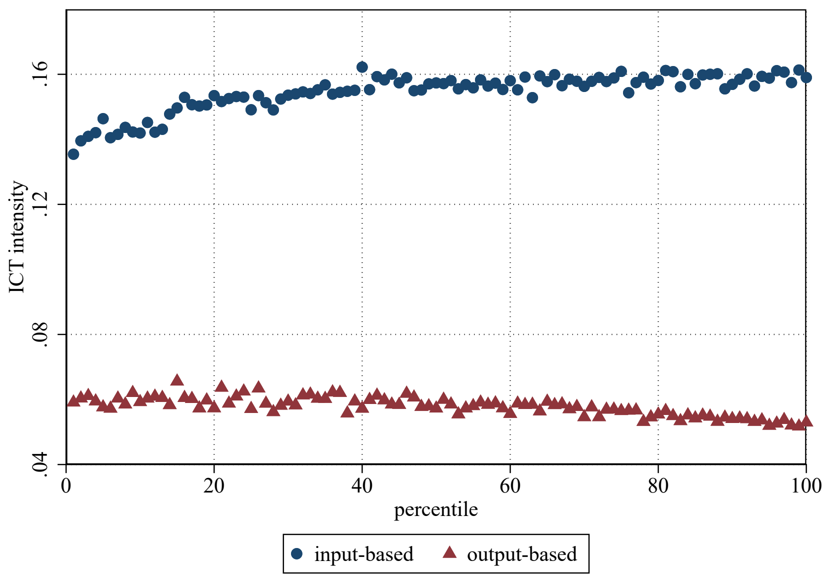
In the main text, we define digitalization based on the inputs used in the production process. A good is classified as digital according to the extent to which it has been produced with digital assets like computer hardware or software. An alternative would be an output-based measure, sorting final consumption goods into digital or non-digital goods. Different attempts have been made to classify final consumption goods. We rely on the BEA's Digital Economy Satellite Account, which represents the BEA's approach to calculating the contribution of digital goods- and service-producing industries to U.S. GDP ("digital economy"). The BEA defines the digital economy as comprising three types of goods and services: digital infrastructure, which primarily refers to ICT, e-commerce and priced digital services, i.e. services related to computing and communication that require a fee (BEA, 2021). The BEA provides data for the digital economy's value added at the industry level. We use the share of digital value-added in total value-added as output-based ICT intensity measure. Since the industry classification is the same as the one we use to construct our output-based measure, we can easily compare it to our input-based measure.

The BEA Digital Economy Satellite Account does not (yet) attempt to estimate the value of free digital goods and services. It is sometimes approximated by the advertising revenue generated by the companies offering it, however, this will insufficiently reflect the consumer surplus generated by these products (Brynjolfsson et al., 2020). Brynjolfsson et al. (2019) and Coyle and Nguyen (2020) present an alternative approach to estimate the value of free digital goods. They carry out choice experiments to estimate consumers' reservation prices, asking them about the monetary compensation at which they are willing to give up access to a digital good such as Google Maps, Facebook or Wikipedia. These studies show that the consumer surplus generated from free goods can be substantial. However, the research in this area is still in its infancy and the existing literature does not study differences between high- and low-income households systematically. Therefore, it is not possible to assess the bias created by ignoring the consumer surplus of free digital goods.

The correlation between our input-based and the BEA's output-based measures is 0.67. While both measures rely highly on the ICT sector (Naics 334* (excl. 3345), 5112, 51913, 5415 and 518), the following categories are highly ICT-intensive according the input-based measure, but not the output-based measure: navigational, measuring, electromedical and control instruments (3345), insurance carriers (5241), newspaper and book publishers (511), accounting and bookkeeping services (5412), other professional, scientific and technical services (5419), administrative and support services (561). These tend to be consumed proportionately more by high-income households. Categories with a low input ICT-intensity but high output ICT-intensity are in particular radio and television broadcasting (5151) and telecommunications (517). The expenditure share of these categories is higher for low-income households. Due to these differences, the ICT intensity along the income distribution looks different when using the output-based definition. Figure A7 shows the digital share of the consumption basket along the income distribution, comparing the two measures. We do not show different sub-periods as the BEA digital economy estimates are available only from 2005. According to the output-based measure, the digital share is much lower for all households, and there are smaller differences between high- and low-income households. Richer households have a slightly lower ICT share now.

This underlines that it matters how digital goods are defined. Choosing one measure or the other might lead to different implications about who benefits and who loses from digitalization. We have opted for an input-based measure in this paper since we are interested in comparing income and price effects of changes in production processes, but we acknowledge that there could also be benefits in studying output-based measures, and that the findings would be complementary to ours.

Figure A7: ICT intensity along the income distribution, output-based classification



Note: The graph shows the ICT share of the consumption basket by percentile for 2005-2017 period comparing the input-based series of the main text with an output-based series constructed based on the BEA's Digital Economy Satellite Account. *Source:* BEA, CEX and own calculations.

A.5 Inflation rate and ICT intensity: Extensions and Robustness

Table A2 shows different regressions of the inflation rate on ICT intensity. Column (1) is our baseline, which we refer to in the main text. Here we regress the average annual inflation rate between 1997 and 2017 on ICT intensity in 2017 at the level of final commodities. In column (2), we repeat the regression but cut the sample as 2012 in order to make the results comparable to

Table A2: Average annual inflation rate and ICT intensity

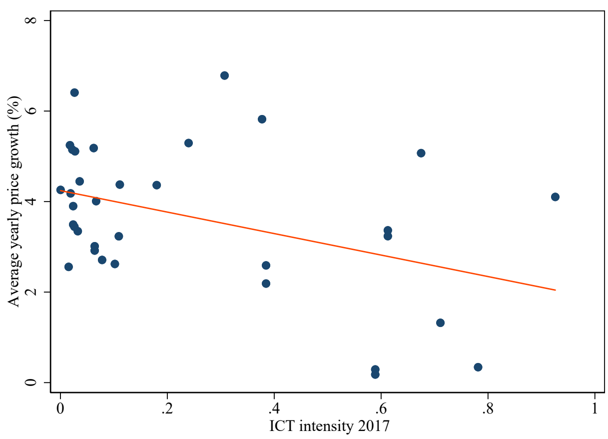
Variables	(1) π_{2017}	(2) π_{2012}	(3) π_{2012}
ICT intensity	-4.220***	-6.136***	-5.152***
	(1.476)	(1.786)	(1.702)
Trade share			-2.835***
			(0.497)
Constant	1.273***	1.422***	2.156***
	(0.298)	(0.333)	(0.341)
Observations	300	281	281
R-squared	0.017	0.037	0.209

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The table shows the regression of average annual inflation rate on ICT intensity. The first column reports the coefficients for data until 2017, the second for data until 2012, while the last includes the import share of industry's overall value added in 2012. *Sources:* BEA, CEX, BLS and own calculations.

column (3), in which we add the import share, calculated as the share of imports in total value added by IO commodity. At the detailed level, the most recent IO tables are for the year 2012. The results show a significant and robust negative association between inflation and the ICT intensity.

Figure A8: Consumer price inflation 1997-2017 and ICT intensity, BEA industries



Note: The graph plots the ICT intensity against average annual inflation in consumer prices between 1997 and 2017. The red line shows the prediction from a linear regression. *Source:* BEA, NIPA and own calculations.

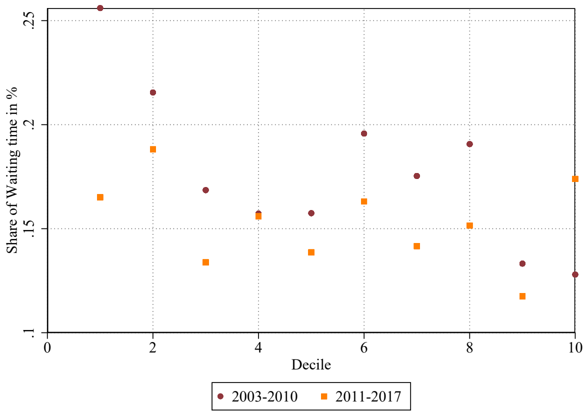
A.6 Consumption and Time Use: Extensions and Robustness

Digital technology could have a large impact on waiting time and search activities in general. Overall waiting time decreases with income. It also has slightly decreased over time, see Figure A9 (a). The time used for shopping of non-groceries that also includes researching of purchases increases with income. It has overall decreased over the last ten years for all income groups, see

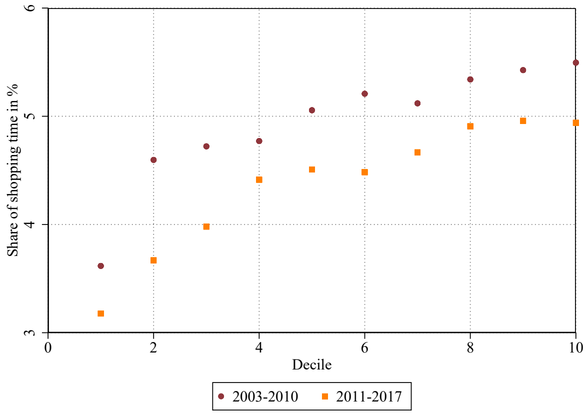
Figure A9 (b). Job search time share decreases the richer a household is. In the years from 2011-1017, the overall share has slightly increased, see Figure A9 (c). We do not know how exactly digitalization affects these dimensions, but it can be expected that it reduces time use. Then, poor households could potentially benefit more regarding job search, whereas rich households could benefit more in terms of shopping time. As all of this is speculative, we refer a detailed analysis to future work.

Figure A9: Time spent (in %) for a category along income

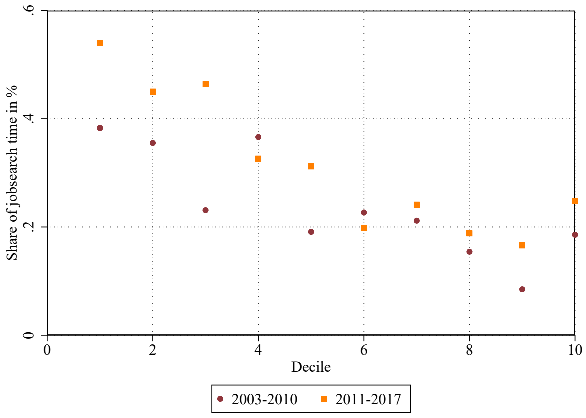
(a) Waiting time



(b) Shopping time



(c) Job search time



Note: The graph shows the share of time that people spent (a) waiting for all activities except personal care, (b) shopping for all activities except groceries (c) searching for jobs. *Source:* ATUS and own calculations.

Table A3: Regression of ICT time use intensity (in %) on household income and age

	Households	
	(1)	(2)
income	0.034***	0.039***
	(0.002)	(0.002)
age		0.042***
		(0.001)
Constant	12.569***	14.198***
	(0.040)	(0.049)
Year FE	Yes	Yes
Observations	142,642	142,642
R-squared	0.069	0.089

Note: Income is in thousands of constant 2015 USD. ICT intensity is in %. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. A low R-squared in the regression of ICT time use and income originates from individual households only reporting income bins. *Source*: ATUS, BEA and own calculations.

A.7 Derivation Compensatory Variation

Suppose an agent gets utility from consuming goods C_i $i \in 1, 2, ..., I$

$$U\left(\{C_i\}_i^N\right)$$

The agent has income m and faces prices $\{p_i\}_i^I$.

Suppose prices change $\{dp_i\}_i^I$. The compensatory variation then asks how much income the price change is worth to ensure the same utility level as before

$$V\left(m + CV, \{p_i\}_i^N\right) = V\left(m, \{p_i + dp_i\}_i^N\right)$$

which is approximately

$$V\left(m, \{p_i\}_i^N\right) + \frac{dV}{dm}CV = V\left(m, \{p_i\}_i^N\right) + \frac{dV}{dp_i}\Delta p_i$$

Recall Roy's identity that gives us an expression for demand x_i of good C_i:

$$-x_i = \frac{dV/dp_i}{dV/dm}$$

Plug this into the approximation above:

$$CV = -x_i \Delta p_i = -x_i p_i \frac{\Delta p_i}{p_i}$$

which states that CV is equal to nominal expenditures times the percentage change of prices. Logic extends when all prices change

$$CV = -\sum_i p_i x_i \frac{\Delta p_i}{p_i}$$

Percentage price changes and the overall amount of expenditures on a certain good category are therefore all we need to compute the equivalent variation in the data. Note that the minus sign in front implies that when prices rise, one needs to deduct income to generate the same loss in utility.

A.8 Income and Asset Data from the SCF

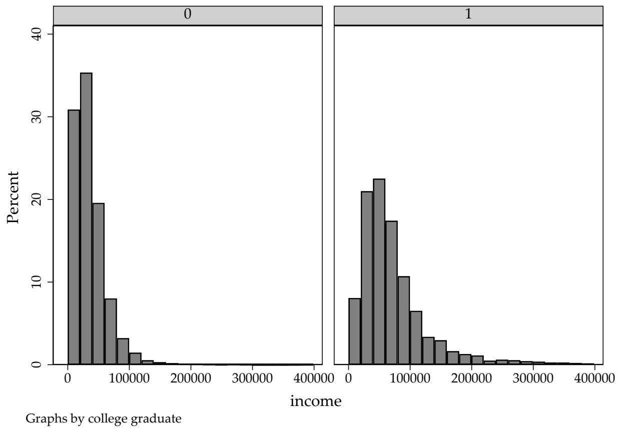
The aim of this section is to motivate our assumption that capital is owned only by high-skill households. To show that this assumption is supported by data, we use data from the Survey of Consumer Finances (SCF) between 1998 and 2013. We focus on households where the head is between 20 and 64 years old and in the labor force. We define incomes as in Kuhn and Ríos-Rull (2016): labor income is the wage and salary income plus a share of business and farm income. This share corresponds to the share of unambiguous labor income (i.e. wage and salary income) in the sum of unambiguous capital income (interest, dividends and capital gains) and labor income. Capital income is interest income, dividends and capital gains plus the remaining share of business and farm income. Total income is the sum of labor income, capital income and transfer income (e.g. social security). It is approximately equal to the payments to the factors of production owned by the household plus transfers, with the exception that it does not include income imputed from owner-occupied housing. Wealth is the household net worth, i.e. financial and non-financial income minus debt. All variables are before taxes.

Table A4: Income and wealth by education group, SCF

	Labor income	Capital income	Total income	Net wealth
college graduates	56,443	50	59,300	142,044
non-college graduates	27,873	0	29,707	27,634

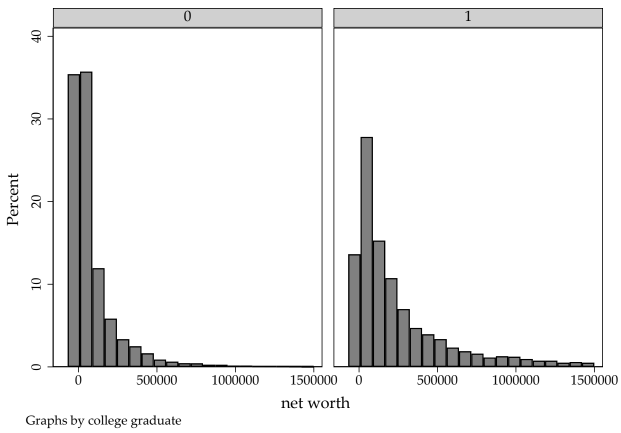
All numbers are in 2019 US Dollars and refer to the median household in each education group. Total income is the sum of labor income, capital income and transfer income.

Figure A10: Income distribution by skill



Note: The graph shows the distribution of income for incomes lower than 400,000 USD. The left panel uses data for households without college degree. The right panel is for college-educated households.
Source: SCF.

Figure A11: Net wealth distribution by skill



Note: The graph shows the distribution of net worth for net worth lower than 1.5 mio USD. The left panel uses data for households without college degree. The right panel is for college-educated households.
Source: SCF.

B Model Appendix

B.1 Solution Algorithm

We solve the model in the following way:

1. Global solution: We solve the model for 10,000 random draws from a holton set of the nine parameter values η, ϵ, ϕ_1, ϕ_2, γ_1, γ_2, ν, θ, ρ within permissible bounds. We define the loss function as the weighted sum of squared errors in the set of data moments (scaled by the corresponding moment). This means that we minimize the percentage error of moments. We choose a weighting matrix that overweights by a factor of five the following, more relevant, moments: the relative price, the skill premium, the wage premium of routine relative to manual workers, the ICT intensity and expenditure share differences between the top 10

and bottom 10 percent. The results are robust to choosing slightly different weights.

2. Local solution: We solve the model for 700 random draws in the vicinity of the 10 best solutions, obtain a new set of 10 best solutions and repeat this procedure three times. From the resulting final set of best solutions, we select the one where the loss function is smallest.

B.2 Link between Parameters and Moments

Several moments link directly to one or more parameters. The difference in the expenditure shares of the ICT-intensive good between two income groups is used as a moment to inform ρ. From eq.(19),

$$\frac{P_1 C_{1M}}{e_M} / \frac{P_1 C_{1H}}{e_H} = \left(\frac{e_H}{e_M}\right)^{\rho} \frac{\Delta_M}{\Delta_H}$$

where we use as empirical approximation the expenditure shares of the top and bottom 10% of the income distribution as derived in Section 2.

θ is pinned down by the increase in the expenditure share of ICT-intensive goods over time. At any t, s,

$$\frac{P_{1t} C_{1t}}{e_t} / \frac{P_{1s} C_{1s}}{e_s} = \frac{\left(\frac{P_{2t}}{e_t}\right)^{\rho} \left(\frac{P_{1t}}{P_{2t}}\right)^{\theta} \Delta_t}{\left(\frac{P_{2s}}{e_s}\right)^{\rho} \left(\frac{P_{1s}}{P_{2s}}\right)^{\theta} \Delta_s}$$

We set $t = 1996$, $s = 2017$ to exploit the change in the expenditure share in our dataset over the full sample period.

The relative quantities consumed of good 1 and good 2 are informative for all three parameters of the utility function:

$$\frac{C_1}{C_2} = \frac{\nu \frac{e_o}{P_1} \left(\frac{P_2}{e_o}\right)^{\rho} \left(\frac{P_1}{P_2}\right)^{\theta}}{\frac{e_o}{P_2} \left(1 - \nu \left(\frac{P_2}{e_o}\right)^{\rho} \left(\frac{P_1}{P_2}\right)^{\theta}\right)}$$

To pin down the parameters of the production function, we choose moments that are informative about the relative use of the inputs R, M, H, ICT and K (which can move freely across sectors).

The skill premium at t of H over M is linked to ϵ,

$$\frac{W_{Ht}}{W_{Mt}} = \frac{\gamma_i (1 - \psi_i) M_{it}}{\psi_i SW_{it}^{\frac{\epsilon-1}{\epsilon}} H_{it}^{\frac{1}{\epsilon}}}$$

Using log differences to proxy for percent changes over time, only M_{it}, SW_{it} and H_{it} are time-varying. γ_i and ψ_i drop out. This leaves ϵ as a parameter to match this moment.

The wage premium of R over M additionally incorporates information about η :

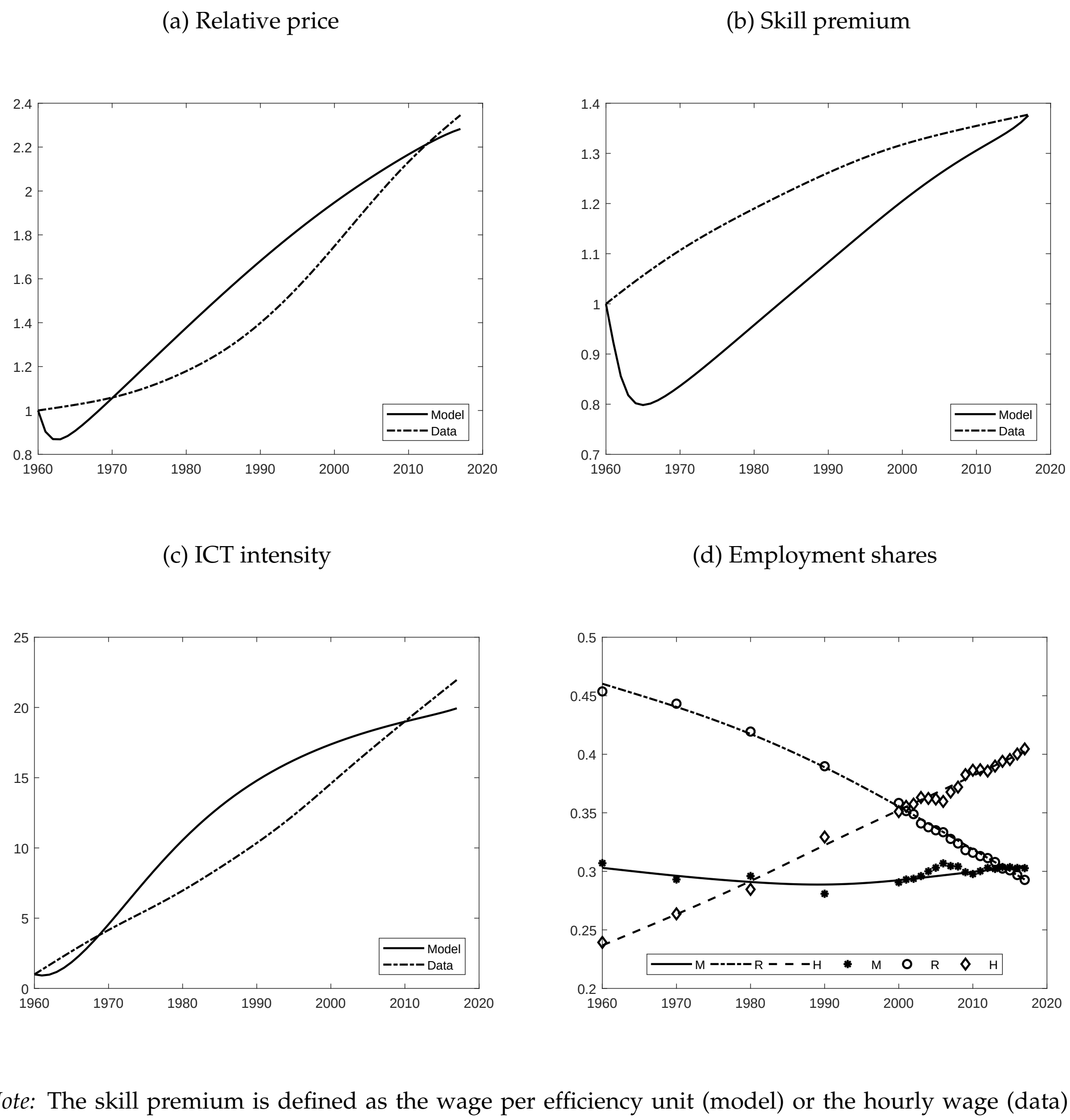
$$\frac{W_{Rt}}{W_{Mt}} = \frac{AW_{it}^{\frac{\epsilon-1}{\epsilon}} (1 - \gamma_i)(1 - \psi_i) \phi_i M_{it}}{\psi_i R_{it}^{\frac{1}{\eta}} AW_{it}^{\frac{\eta-1}{\eta}} SW_{it}^{\frac{\epsilon-1}{\epsilon}}}$$

Considering log differences, the two parameters that remain to match this moment are ϵ and η.The time series for $W_{Rt/}W_{Mt}$ is used to match the employment share of routine and manual workers over time.

To match input factors and their returns, we choose the following additional moments: the ICT intensity in the aggregate economy, the return on ICT capital and the labor share, which is informative about the use of capital vs. labor in the production process. Koh et al. (2020) show that the decline in the labor share is fully explained by the rise in intellectual property products (whose rents are attributed to capital income by the BEA). A major share of intellectual property products is software. Karabarbounis and Neiman (2014), Martinez (2021), Eden and Gaggl (2018) and Acemoglu and Restrepo (2018) also argue that automation is an important driver of the labor share. Hémous and Olsen (2022) choose a similar calibration strategy to ours where they also require that automation explains the declining labor share and the increasing skill premium fully. Finally we match the evolution of the relative price from Figure 9. This is decisive for the overall strength of the price effect and also helps to inform other parameters in the production function.

B.3 Additional simulation results

Figure B.1: Calibration fit for the most important moments



Note: The skill premium is defined as the wage per efficiency unit (model) or the hourly wage (data) of high-skilled workers relative to routine workers.