Consumption Inequality in the Digital Age*

Kai Arvai†
Banque de France

Katja Mann‡
Copenhagen Business School

13 June 2022

Abstract

This paper studies how digitalization affects consumption inequality. We assemble a novel dataset of digital technology used in the production process, link it to US consumption data and establish a new stylized fact: High-income households consume a higher share of digitally produced products than low-income households. Building on this finding, we present a structural model in which digitalization affects consumption inequality in two ways: By a polarization of incomes and by a decline in the relative price of digitally produced goods. Both channels work in favor of high-income households. Calibrating the model to the US economy between 1960 and 2017, we demonstrate that the price channel has sizeable welfare effects and amplifies the increase in consumption inequality that is caused by digitalization by around 25%.

Keywords: Digitalization, Inequality, Consumption, Income

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

*We would like to thank Christian Bayer, 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 consume and produce. The increased usage of computers and other digital assets in production has contributed to 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] High-skilled workers have become more productive, whereas low-skilled or routine-intensive jobs have declined. As a result, the skill premium has increased, contributing to a rise in income inequality. However, progress in digital technology does not only affect wages: Sectors that rely more on digital capital in their production process can produce at a lower cost and offer lower prices to their customers. As consumption patterns differ across the income distribution, these price changes will affect rich and poor 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 in the United States.

By combining household- and sector-level data, we establish that high-income households are the main beneficiaries of digitally-induced price changes, as they consume goods and services that are produced with a higher intensity of digital capital. Income and price changes thus work in the same direction, making the distribution of consumption more unequal. To assess how important the price changes are quantitatively, we present a structural growth model, which we calibrate to the US economy between 1960 and 2017. When explicitly including price changes, the effect of digitalization on consumption inequality is around 25% larger than what the income effect alone would suggest.

The focus of this paper is on input-based digitalization, by which we understand the degree to which industries use digital assets in the production process.[2] *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. Our resulting digitalization measure covers 61 industries between 1960 and 2017. We document that the share of ICT capital in the overall capital stock has increased from almost 0% in 1960 to over 16% in 2017. Importantly, there is substantial 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 digitalization content of intermediate products by relying on the input-output structure of the production network.

[1] When considering automation more broadly, the empirical evidence is even ampler, see e.g. Autor et al. (2003), Autor and Dorn (2013), Cortes et al. (2017), Eden and Gaggl (2018) and Hémous and Olsen (2022). Prettner and Strulik (2020) and Moll et al. (2021) show that automation also increases wealth inequality.

[2] Alternatively, goods could be classified based on whether they themselves constitute digital goods. Appendix A.3 discusses this issue in more detail. The input-based measure is better suited to compare how changes in production technology affects incomes and prices.

In the next step, we link this digitalization measure of final commodities to consumption categories in the Consumption Expenditure Survey (CEX). Using CEX data on household income, 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. We also document that consumer price inflation has been weaker for ICT-intensive commodities.[3]. Furthermore, we perform a back-of-the-envelope calculation that considers the difference in ICT intensity of consumption baskets jointly with the relationship between price changes and ICT intensity. It reveals that the price index between different income groups diverges by several percentage points over the decades. In a compensatory variation, we show that the poorest households need to be compensated by about 40% of their initial income to be indifferent to the price increase of their consumption basket, whereas the rich require only 20%.

In the second part of the paper, we assess the strength of the price effect relative to the ICT-induced income changes. To this end, we introduce a structural model building on our empirical findings. The model features a two-sector economy with two types of capital, ICT and non-ICT. Sector 2 uses ICT capital more intensively than sector 1. The economy is populated by two types of agents that differ by skill endowment. Following Eden and Gaggl (2018), high-skill labor and ICT capital produce a composite good, which is combined with low-skill labor. The resulting composite is then used together with non-ICT capital to produce the final good. Digitalization is modeled as an increase in the rate of transformation of output into ICT capital. We introduce non-homothetic preferences as in Boppart (2014), which imply that the effect of changing relative prices depends on the agent's position in the income distribution. We estimate the preference parameters directly from our newly assembled data and confirm that ICT-intensive goods are luxury goods. We also find that the elasticity of substitution between the ICT-intensive and non-intensive goods is slightly above 1, making them partial substitutes. To our knowledge, we are the first to provide parameter estimates for household preferences over ICT- and non-ICT goods.

We calibrate the other parameters of 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 and the decline in the price of ICT-intensive goods. Low-skill labor needs to be a substitute to ICT capital and high-skill labor. This generates the increase in wage polarization we observe in the data. The simulated increase in income inequality goes along with a decline in the price of ICT-intensive goods, benefiting the high earners as well. In an analysis of the overall welfare effects, we show that high-skill households experience an increase in welfare that is equivalent to 28.5% of their initial income, whereas low-skill households experience an increase of only 2%. As the price indices of both households diverge, consumption inequality increases by 25% more than the increase in income inequality alone suggests. The dynamic of the model-based price indices are in line with the back-of-the-envelope calculations in the empirical part. These results imply that the price channel is quantitatively important.

There exists ample evidence documenting an increase in consumption inequality over the last decades (see Attanasio and Pistaferri (2016) and references therein). To our knowledge we are the first to consider consumption inequality in the context of digitalization. The topic has received a

[3]This aligns with the more general finding that inflation is higher for the poor, e.g. Kaplan and Schulhofer-Wohl (2017) and Jaravel (2019); see Jaravel (2021) for an overview.

broader interest by the trade literature, e.g. Fajgelbaum and Khandelwal (2016), Nigai (2016) and Borusyak and Jaravel (2021). The effect of digitalization is similar and different from that of trade: Similar because in both cases, a change in the production process alters factor returns as well as the price of output, and different because automation works through the capital stock and thus attributes a crucial role to the complementarity between capital and labor. Also related to this paper is Hubmer (2018), who studies the labor intensity of consumption baskets along the income distribution and shows that high-income households spend relatively more on labor-intensive goods and services.

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

This section motivates our focus on the price channel. We assemble a novel dataset to establish that households along the income distribution differ in the consumption share of digitally produced goods. Our data show that the rich consume more ICT-intensive goods than the poor. At the same time we find that prices of ICT-intensive goods have grown at a slower pace than prices of non-ICT goods, making the rich households the prime beneficiaries of digitalization. We use a compensatory variation to provide a first estimate of the welfare effects of digitalization along the income distribution.

2.1 Measuring the ICT-intensity of consumption

We assemble our dataset by focusing on industry-level capital stocks and household-level consumption and proceed along the following steps: We first create an industry-level measure of ICT intensity by computing ICT vs. non-ICT capital used in production. 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 digitalization share of consumption baskets along the income distribution.

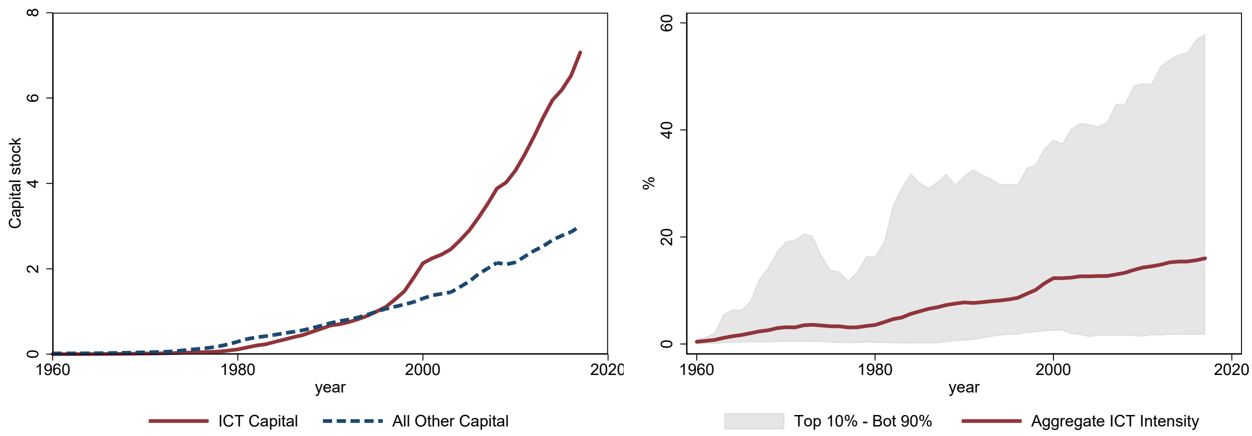
2.1.1 ICT capital share by industry

The BEA provides data on the stock and investment of 96 different types of capital for 61 private industries in the Detailed Data for Fixed Assets. We use this yearly dataset between 1960-2017 to construct the stock of digital capital for each industry. Capital is evaluated at its current price. 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, both stocks indexed to 1 in 1995. Over the whole time horizon, ICT capital has grown faster than non-ICT capital. Since 1995, the ICT capital stock has increased by more than 700%, while non-ICT capital has increased by around 300%.

Figure 1: ICT capital

(a) Aggregate capital stock (1995=1) (b) Industry-level ICT intensity



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.

In the next step, we construct a measure of the importance of digital capital in the industries' production process. The ideal measure would be the share of expenses for digital capital in overall expenses. In the data however, capital expenses (the factor payments $R \cdot K$) for digital assets are not directly observable, as firms usually own their assets. There exist estimates for the rate of return for ICT capital, but these vary depending on the estimation methodology. Compare for example the measures proposed by Eden and Gaggl (2018) and Karabarbounis and Neiman (2014) in Figure A9 in the Appendix or Barkai (2020) (at the industry level). In contrast, the stock of capital and its components are directly observable. Throughout the paper, we will therefore measure the degree of digitalization as the share of ICT capital in the overall capital stock of an industry or of the whole economy. We refer to this measure as **ICT intensity**. Alternatively as a robustness check, we also include the overall labor share on value added of an industry: This measure is the share of digital capital on the capital stock multiplied with the overall capital share (one minus the labor share)[5]. Note that, 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-

[4]This measure misses digital assets that are being offered for free. Software or digital services that do not have a market price will not show up as investment in the statistics. But while digital products are often offered for free to private consumers, they usually need to be purchased when used for business purposes, e.g. e-mail management systems, antivirus software, file hosting services, etc.

[5]Figure A5 shows this measure in the Appendix. High ICT-intensive industries do not exhibit systematically different patterns in the labor share. As the labor share is only imperfectly measured (see Elsby et al. (2013)) we prefer the ICT capital intensity measure.

and non-ICT technology.

Ignoring any interlinkages between industries, the right panel of Figure 1 shows the ICT intensity over time. We focus on 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. 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 (2021) 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 Hubmer (2018) or Aghion et al. (2020).

2.1.2 Digitalization share of final output

Industries may not only use digital capital in their own production, 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 Accounts show how industries provide input to or use output from each other. These accounts provide detailed information on the flows of the goods and services that comprise the production process of industries. We use the detailed Input-Output tables after re-definitions and focus on private industries. 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 to calculate the digitalization share of final goods and services as a weighted sum of the digitalization shares of its intermediate inputs and value added. See Appendix A.1 for details.

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 sum of the digitalization shares of all inputs plus 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. We construct the ICT-intensity measure for 1996-2017 for all IO commodities. This time period is chosen because we will later link the industry data to the CEX, which is available only from 1996. Commodities that

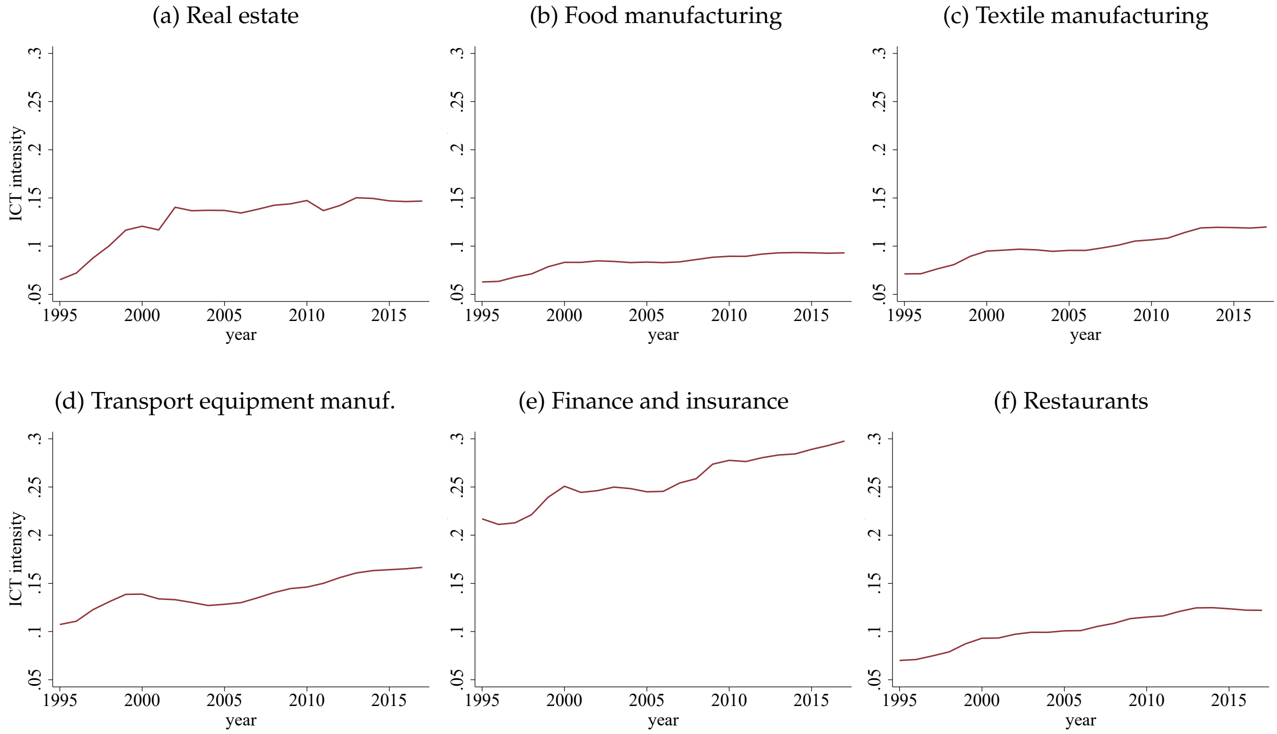
[6] The temporary drop for the top 10 % in the 1970s is due to declining ICT investment activity and high depreciation rates in the sector for credit intermediation and in scientific research.

[7] All sectors except for the construction sector are matched. For 1996-2004 we use the 1997 matrix for 2005-2017 we use the 2012 matrix.

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 (see Figure A1).

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.[8] 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 value of ICT intensity and notable finance and insurance has a much higher ICT intensity throughout the whole sample period whereas food manufacturing has a low ICT intensity. These differences will become relevant later on.

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.

2.1.3 Consumption patterns across households

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 information on the households. This enables

[8]The following IO commodities are in these categories: real estate: 5310HS; food manufacturing: 311*, 312*; textile manufacturing: 313*, 314*, transport: 336*; finance and insurance: 52*; restaurants: 722*.

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.[9]

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. [10]

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

industry	1996-98			2016-18		
	1st	10th	**ratio**	1st	10th	**ratio**
real estate (incl. construction)	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**

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 1996-98 and for 2016-2018. These are ranked by their importance for low-income households. The largest share of income is being spent on housing and therefore

[9]While the CEX is the dataset best suited for our analysis and a standard reference in the literature on consumption inequality (e.g. Aguiar and Bils, 2015, Attanasio and Pistaferri, 2016), it nevertheless has certain shortcomings, which we discuss in Appendix A.1.

[10]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.

gets allocated to real estate. There has been some increase 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[11]. 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 in most categories. Notable exceptions are education, whose share in the expenditure basket of the rich more than doubled while dropping for the poor, and professional services, whose importance increased for the rich while dropping for the poor.

2.1.4 Digitalization share of consumption

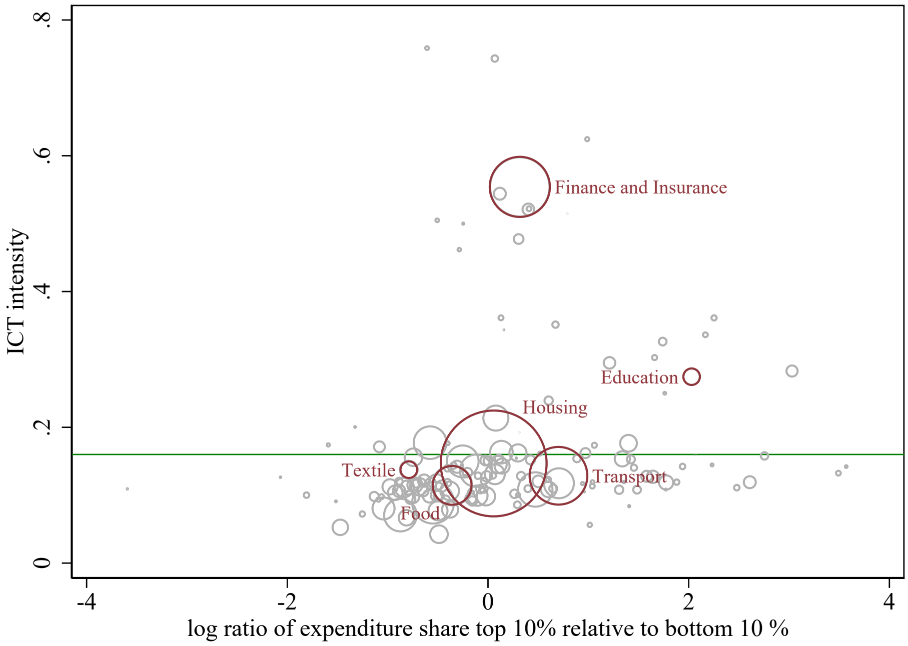
Before considering the ICT intensity of different consumption baskets as a whole, we illustrate which commodities are particularly relevant for determining the ICT intensity of high- vs. low-income consumption. In Figure 3, we plot the ICT intensity against the relative importance of a commodity (in terms of expenditure share) for high-income vs. low-income households. We consider data for final commodities in the year 2017[12]. 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) in 2017.

There is a positive correlation between the ICT intensity and the relative expenditure share of the rich. On the left side of the graph, where categories are more important for poor households, most ICT intensities are below the average value of 0.16. These are in particular goods produced by (non-ICT) capital-intensive manufacturing industries like textiles or food. Around the center, where commodities are equally important for both households, 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 that are relatively more important for the rich, we have some categories of average ICT intensities, but also some with very large ICT intensities such as education or health.

[11] Finance and insurance may be considered an investment good as well as a consumption good. To show that the results in Section 2.1.4 are robust to excluding this category, we repeat the analysis in Appendix Figure A11 without the IO commodities that relate to finance and insurance.

[12] Data for 1997 and 2007 are provided in the Appendix, Figure A10.

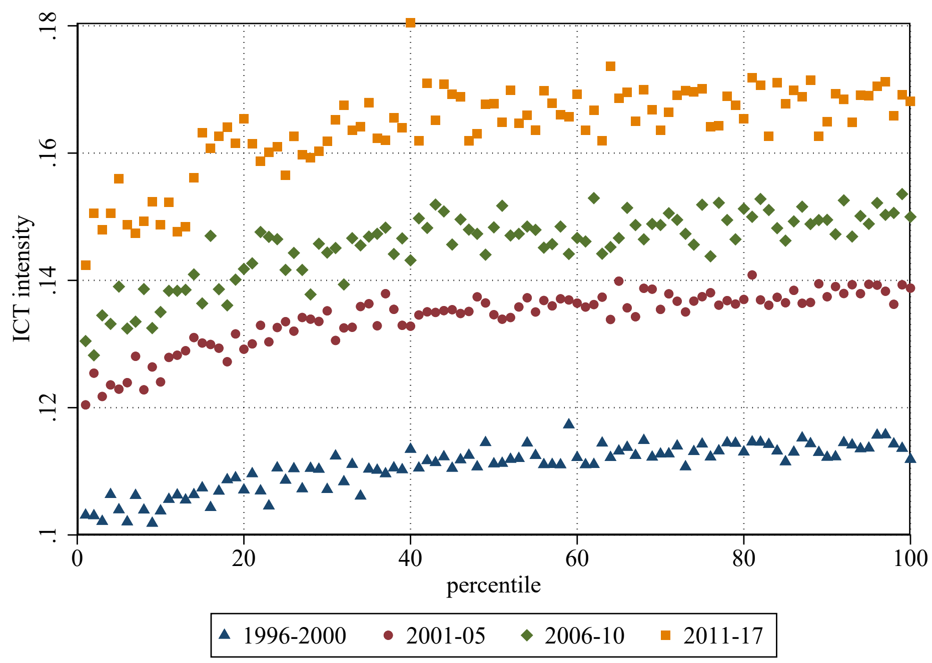
Figure 3: ICT intensity of vs. relative expenditure shares by commodity in 2017



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.

As the final step, we compute the ICT intensity of the 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. So the strong increase in the ICT share over time that we documented in Section 2.1.1 affects the consumption bundle of the rich slightly more. The inequality in digital consumption has thus been increasing over the last 20 years.

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.

Is the different ICT share for high- and low-income households consequential? In the most recent period, the households at the lower end of the income distribution have an ICT share of 15% in their consumption bundle, while those at the upper end have an ICT share of around 17%. Considering lifetime consumption, these numbers could potentially have strong welfare effects of digitalization.

2.2 An empirical estimate of the price effect

In the following, we provide a first estimate of the welfare effects of digitalization based on our dataset. We document large differences in consumer price inflation across goods and investigate to which extent the different ICT intensities in the consumption baskets of richer and poorer households are generating heterogeneous dynamics of inflation. We then carry out a compensatory variation which quantifies the welfare loss of inflation along the income distribution.

2.2.1 Digitalization and consumer prices

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.[13] We match by hand 632 CEX product categories with 318 BLS price data series between 1996 and 2017,

[13]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.

similar to Jaravel (2019).[14] We then create Törnqvist price indices for each income decile. The Törnqvist index is a chain-weighted price index that considers product substitutions made by consumers and other changes in their 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 the median household s_i,

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

The level can be recovered recursively relative to a base year 0 with $p_0 = 1$ as $p_t = p_{t-1}\exp(\Delta\pi_t)$. Columns (1) and (2) of Table 2 show the evolution of the annual inflation rate for households in the 1st and 10th income decile. Both groups have experienced a similar price dynamic over time. An exception is the period between 2004-2011. In this period, the poorest 1st decile experiences higher inflation rates. This is in line with Jaravel (2019) and Kaplan and Schulhofer-Wohl (2017), who find that inflation has been higher for low-income households than for rich households in this period, see also Figure A12 in the Appendix. As in Hochmuth et al. (2022), the trend begins to reverse at the end of our sample.

To understand the relevance of ICT-intensive goods for the dynamic of expenditures shares, it is helpful to consider an ICT inflation price index. For this purpose, we focus on the most ICT-intensive goods (goods with an ICT intensity of more than 40% in 2017; these comprise 10% of total expenditures of the median household) and compute inflation for this consumption bundle.

Columns (3) and (4) of Table 2 show the inflation rate for ICT-intensive goods. Remarkably, the ICT inflation rate has been lower for the richest 10th decile than for the poorest 1st decile and helped to keep the inflation rate of the overall basket for the rich lower, while ICT inflation was often higher for the poorest decile. A key role is played by the expenditure shares. To show this, we construct an analogous inflation index for ICT-intensive goods in column (5) and (6), but compute annualized inflation with fixed expenditure weights of 1996 and 1997.

To infer whether households have shifted their expenditures away from goods that experienced higher inflation, we consider the difference between the actual ICT index and the index with fixed weights in columns (7) and (8).

The difference is almost always negative for the richest 10th decile, meaning that the expenditure share of low-inflation goods has risen within the ICT bundle. Households have shifted their demand towards these goods, suggesting some degree of substitutability for the rich. The picture for the first decile is less clear, as the difference is negative in the first half of our sample, but positive thereafter. We conduct a corresponding decomposition for the overall basket in Appendix Table A2. There, we demonstrate that the overall basket is characterized by complementarity, as the expenditure share of high-inflation goods has increased for all income classes.

Next, we correlate ICT-intensity of each consumption category with its average yearly price change. As Figure 5 shows, consumer price inflation has been lower between 1997 and 2017 for consump-

[14] 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.

Table 2: ICT Inflation decomposition for 1st and 10th decile

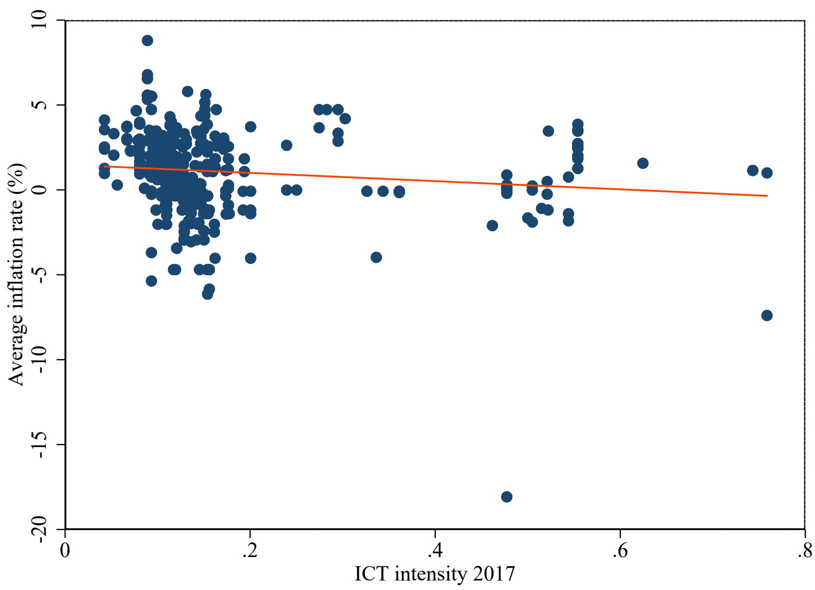
year	π_t		π_t^{ICT}		$\pi_t^{ICT,fixed}$		Δ^{ICT}	
	1st	10th	1st	10th	1st	10th	1st	10th
1996	-	-	-	-	-	-	-	-
1997	0.9	1	2.2	2	2.2	2	0	0
1998	1	1.2	0.8	0.9	0.8	0.9	0	0
1999	2.6	2.6	1	0.8	1	0.8	0	0
2000	3.1	3	2	1.5	1.9	1.7	0.1	-0.2
2001	0.9	1	4.9	4.4	5	4.5	-0.1	-0.1
2002	2.2	2.3	5.5	4.8	5.7	4.8	-0.2	0
2003	1.3	1.4	2.7	2.4	2.8	2.2	-0.1	0.2
2004	3.4	3.1	2.8	2.5	2.6	2.2	0.2	0.3
2005	3.3	2.9	1.5	1.4	1.6	1.5	-0.1	-0.1
2006	2.2	2.4	0.4	-0.2	1.1	1.1	-0.7	-1.3
2007	4.6	4	0.4	-0.3	1.2	1.5	-0.8	-1.8
2008	-1.4	-0.5	-0.7	0.7	0.9	3	-1.6	-2.3
2009	3.6	3.2	0.9	-0.4	0.5	2.5	0.4	-2.9
2010	1.7	1.5	1.8	0.6	0.4	2.2	1.4	-1.6
2011	3.5	3.1	1.6	1.2	0.9	2.3	0.7	-1.1
2012	1.6	1.8	2.4	1.5	0.8	2.9	1.6	-1.4
2013	1.3	1.4	1.6	1	0.4	2.1	1.2	-1.1
2014	0.1	0.4	1.6	1.4	0.2	2.4	1.4	-1
2015	-0.3	0.4	2.6	2.5	0.5	3.2	2.1	-0.7
2016	1.3	1.8	2.6	3	0.7	3.7	1.9	-0.7
2017	1.9	1.7	2.1	0.6	-0.2	3	2.3	-2.4

Note: The table shows the decomposition of inflation for the first and the tenth income decile. Inflation π_t is the inflation rate of the overall index, π_t^{ICT} is the index for the most ICT-intensive goods with changing weights and $\pi_t^{ICT,fixed}$ is inflation with fixed 1996/1997 weights for the ICT-intensive basket. ICT-intensive goods are those with $> 40\%$ ICT in 2017. Rounding errors explain residuals in the decomposition. *Source:* BEA, CEX and own calculations.

tion categories that have a higher ICT intensity. Among the categories with the largest price decline are applications, games and televisions[15]. The largest price increases are in low-ICT intensity categories such as fuels, gasoline and medical services, where obviously, inflation has been driven by factors not related to technological change. A linear regression of price changes on ICT intensity estimates a slope coefficient of -2.4, which is highly significant (standard deviation: 0.79; see Appendix Table A1 for the full regression results). This finding aligns 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 also leads to lower prices. As a robustness check, in Appendix A.4, we also include the import share, 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. Indeed, a higher import share is associated with lower price increases. When including the import share in the regression, the coefficient 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 the Appendix A.4. The correlation between ICT intensity and inflation is also negative and significant (Appendix Figure

[15] Televisions constitutes a big outlier with -17% average yearly inflation and an ICT intensity of 48% in 2017.

Figure 5: Consumer price inflation 1997-2017 and ICT intensity



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, CEX, BLS and own calculations.

A8). Finally, we conduct a back-of-the-envelope calculation in which we combine the difference in ICT intensity of two percentage points between the bottom and the top with the slope coefficient of the regression of average price changes on ICT intensity (-2.4): Over the time period between 1960 and 2017 the different price dynamic sums to a 2.8% difference in the price index for the top and the bottom income group[16].

2.2.2 Compensatory variation using the data

Combining our findings on differences in the ICT intensity of consumption baskets (Section 2.1.4) with the information on price changes in the previous section, we provide 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 carry out a compensatory variation. It tells us how much additional income we need to provide to households to make them indifferent to the price increases that took place between 1997 and 2017. We can approximate this by

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

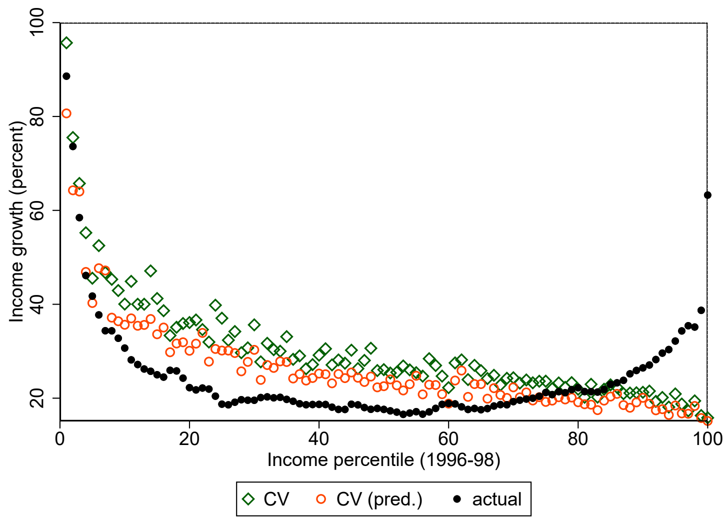
where $x_{ij}p_i$ is household j's expenditure on commodity i in the base year 1997 (price p_i times quantity x_{ij}). $\frac{\Delta p_i}{p_i}$ is the price change between 1997 and 2017. See Appendix A.6 for the derivation. The equation tells us that the money needed to compensate households for price changes equals the initial consumption bundle times the change in prices for each commodity. This formula is a linear approximation and does not feature any changes in the consumption bundle between 1997 and 2017. Households respond neither to changes in relative prices, nor to changes in income. Considering the data in Table 1, expenditure shares across broad product categories have indeed

[16] $(1 + 2.4/100 \cdot 0.02)^{57} = 2.8\%$

remained relatively stable between 1997 and 2017. The decomposition of the overall consumption basket seems also to suggest that substitutability is not the main driver as temporary higher prices for a good translate into temporary higher expenditure shares, in line with the analysis here.

Figure 6 shows CV_j/m_j (green squares), where m_j is the 1997 income, and $\Delta m_j/m_j$, the actual change in nominal income (black circles) between 1997 and 2017. Furthermore we also show the predicted ICT-induced price changes that are based on the regression in the previous section (red circles). This highlights price changes originating from changes in digital technology.

Figure 6: Compensatory variation and actual income changes



Note: This graph plots the compensatory variation (green squares), the actual nominal increase in income (black circles) and the compensatory variation for ICT induced price changes for each income percentile between 1997 and 2017. *Source:* BEA, CEX, BLS and own calculations.

Focus first on the actual increase of nominal income in the CEX dataset. There is clear evidence of income polarization, as documented e.g. by Goos and Manning (2007), Goos et al. (2009) and Autor and Dorn (2013). The upper class and the very bottom of the income distribution experienced substantially larger increases than the middle class. Compare this to the effect of price changes, first considering overall price changes. Our compensatory variation demonstrates that price changes have been more disadvantageous for poorer households. The poorest 10 % of households need to be compensated by on average more than 40% of their initial income for the price increase their consumption baskets experienced. This eats up much of their 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 20% more of their initial income for 20 years of price inflation of their basket. If we focus on predicted price changes due to rising ICT intensity, the compensation required is a bit lower, as not all of the price increase relates to higher ICT intensity. Importantly, the difference between the pure compensatory variation and the one with ICT induced values is greater for poor households. This indicates that ICT-induced price changes are especially important for the rich.

As the back of the envelope calculation entails several simplifying assumptions, the numbers provided in this exercise give only a rough idea of the price effect. Furthermore, this exercise does speak to the income effect of digitalization, whereas it is interesting to compare the effects of ICT-induced price changes with those of ICT-induced income changes. Our structural model in Section 3 addresses these issues. Nevertheless, this exercise suggests that the price effect is quantitatively important.

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, there are two capital goods (ICT and non-ICT) and two sectors with different demands for ICT capital, who each produce a different final consumption good. The model world is populated by two types of households with non-homothetic preferences, which implies that expenditure shares of the two goods depend on total consumption expenditures. 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

There are two sectors $i = 1, 2$, which each produce a final output using as inputs two types of capital and two types of labor. We follow Krusell et al. (2000) and Eden and Gaggl (2018) in setting up a nested production structure that consists of several production functions with constant elasticity of substitution (CES).

Inner nest: High-skill labor H_i and ICT capital ICT_i are used to produce skilled work output (SW):

$$SW_i = \left[\gamma_i H_i^{\frac{\epsilon_i - 1}{\epsilon_i}} + (1-\gamma_i) ICT_i^{\frac{\epsilon_i - 1}{\epsilon_i}}\right]^{\frac{\epsilon_i}{\epsilon_i - 1}}, \tag{3}$$

with ϵ_i being the elasticity of substitution between high-skill labor and ICT capital. Smaller values indicate that both inputs are complements. Values that are larger indicate that they are substitutes. γ_i governs the importance for each input factor. Note that if $\epsilon_i = 1$, the function is Cobb Douglas with factor share γ_i.

Middle nest: Total work output (TW) is produced by combining SW and low-skill work L_i:

$$TW_i = \left[\phi_i L_i^{\frac{\eta_i - 1}{\eta_i}} + (1-\phi_i) SW_i^{\frac{\eta_i - 1}{\eta_i}}\right]^{\frac{\eta_i}{\eta_i - 1}}, \tag{4}$$

with η_i as the elasticity of substitution between low-skill labor and the skilled work composite. If η_i is large, then low skill labor tends to be more substitutable to skilled work. In our specification, there is no direct elasticity of substitution between ICT capital and low-skill labor. Rather it is the composite of ICT and high-skill work that substitutes (or complements) low-skill labor.

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 (Eden and Gaggl, 2018 and 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. They pay out wages W_H and W_L and they rent capital which pays interest rates R_K and R_{ICT}. First-order conditions yield the following equations for factor returns

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

$$L_i: \quad P_i(1-\alpha_i)\phi_i \frac{Y_i}{TW_i^{\frac{\eta_i - 1}{\eta_i}} L_i^{\frac{1}{\eta_i}}} = W_L, \tag{7}$$

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

$$ICT_i: \quad P_i(1-\alpha_i)(1-\phi_i)(1-\gamma_i) \frac{Y_i SW_i^{\frac{\eta_i - 1}{\eta_i}}}{SW_i^{\frac{\epsilon_i - 1}{\epsilon_i}} TW_i^{\frac{\eta_i - 1}{\eta_i}} ICT_i^{\frac{1}{\epsilon_i}}} = R_{ICT}. \tag{9}$$

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 K while 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

$$Y_1 = C_1 + I_K, \qquad Y_2 = C_2 + \mu I_{ICT}. \tag{10}$$

μ is the relative price of ICT capital and at the same time measures 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.[17]

[17] 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.

Capital follows the standard law of motion

$$K' = (1-\delta_K)K + I_K, \qquad ICT' = (1-\delta_{ICT})ICT + I_{ICT}, \tag{11}$$

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

3.3 Households

There are two types of households in this economy, high-skill and low-skill. Think about skill as level of education, which is determined before entering the labor market. Both types $j = H, L$ 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, as originally defined by Muellbauer (1975). 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

$$V(P_1, P_2, e_j) = \frac{1}{\rho}\left(\frac{e_j}{P_2}\right)^{\rho} - \frac{\nu}{\theta}\left(\frac{P_1}{P_2}\right)^{\theta} - \frac{1}{\rho} + \frac{\nu}{\theta}, \tag{12}$$

where the parameters govern the elasticity of substitution between goods (θ), their average expenditure shares (ν) and the income elasticity (ρ). We do not make any explicit parameter restrictions at this point.[18] 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. PIGL preferences also allow for easy aggregation. Finally, they have an explicit empirical justification and are widely used in demand system estimations. The parameters of the indirect utility function eq. (12) can be estimated using the micro data on expenditure shares across households and over time which we assembled in the previous section.

The high-skill households jointly provide $\bar{H}$ units of labor, the low-skill households $\bar{L}$. We keep the labor supply fixed and exogenous. Households can work in any sector and can switch at no cost. All non-ICT and ICT capital is owned by type H households. This assumption is motivated by empirical evidence, as we show based on the Survey of Consumer Finances (SCF) in Appendix A.7.[19] 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 type H households is

$$\underbrace{C_{1,H}P_1 + C_{2,H}P_2}_{e_H} + I_K P_1 + I_{ICT} P_2 = \bar{H} w_H + K r_K + ICT r_{ICT}, \tag{13}$$

[18] 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.

[19] Jaimovich et al. (2020) make the same assumption about capital ownership.

where r_{ICT} is the depreciation-adjusted net interest rate. L households are excluded from capital markets and are hand-to-mouth consumers. The budget constraint for type L households is

$$\underbrace{C_{1,L}P_1 + C_{2,L}P_2}_{e_L} = \bar{L}w_L. \tag{14}$$

Both 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:

$$C_{1,j} = \nu \frac{e_j}{P_1} \left(\frac{P_2}{e_j}\right)^{\rho} \left(\frac{P_1}{P_2}\right)^{\theta}, \qquad C_{2,j} = \frac{e_j}{P_2} \left(1 - \nu \left(\frac{P_2}{e_j}\right)^{\rho} \left(\frac{P_1}{P_2}\right)^{\theta}\right) \tag{15}$$

The elasticity of substitution between goods 1 and 2, σ_j, is household-specific and given by[20]

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

Only the H-type households solve an intertemporal optimization problem, choosing expenditures and investment.

The value function can be expressed as

$$W(K, ICT) = \max V(P_1, P_2, e_H) + \beta \mathbb{E}\left[W(K', ICT')\right], \tag{16}$$

which we solve subject to the household budget constraint eq. (13). Optimization results in the Euler equations

$$\left(\frac{P_2'}{P_2}\right)^{\rho} \left(\frac{e_H'}{e_H}\right)^{1-\rho} = \beta \frac{\mu'(1 - \delta_{ICT}') + R_{ICT}'}{\mu}, \tag{17}$$

$$\left(\frac{P_2'}{P_2}\right)^{\rho} \left(\frac{e_H'}{e_H}\right)^{1-\rho} = \beta \left((1 - \delta_K') + R_K'\right), \tag{18}$$

which we can also write jointly as a no-arbitrage condition

$$(1 - \delta_K') + R_K' = \frac{\mu'(1 - \delta_{ICT}') + R_{ICT}'}{\mu}. \tag{19}$$

This condition states that effective returns on ICT capital have to equal effective returns on non-ICT capital.

[20] See Boppart (2014) for the proof.

3.4 Market clearing

The two capital markets clear

$$K = K_1 + K_2, \qquad ICT = ICT_1 + ICT_2. \tag{20}$$

The markets for consumption goods clear

$$C_1 = C_{1,L} + C_{1,H}, \qquad C_2 = C_{2,L} + C_{2,H}. \tag{21}$$

and finally the labor markets clear

$$\bar{L} = L_1 + L_2, \qquad \bar{H} = H_1 + H_2. \tag{22}$$

4 Calibration and Simulation

This section explains the calibration. We first construct a series of the relative price of ICT capital μ from the data. The change in μ represents our exogenous digitalization shock. Following Boppart (2014), we use the dataset described in Section 2 to directly estimate the preference parameters. By exploiting variation in expenditure shares of ICT-goods across income classes and over time, we identify the parameters that govern the income and substitution elasticities in the utility function. The parameters in the production function are then calibrated to match key moments from the data using the Simulated Method of Moments (McFadden, 1989). We calibrate the model such that digitalization fully explains the changes observed in the skill premium and the relative price in the U.S. economy.

In the simulation, we assume that the economy is at a pre-digitalization steady state in 1960. Between 1960 and 2017, digitalization takes place. μ is the only time-varying parameter. 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 skill type, and how strong the price and income channels are, respectively.

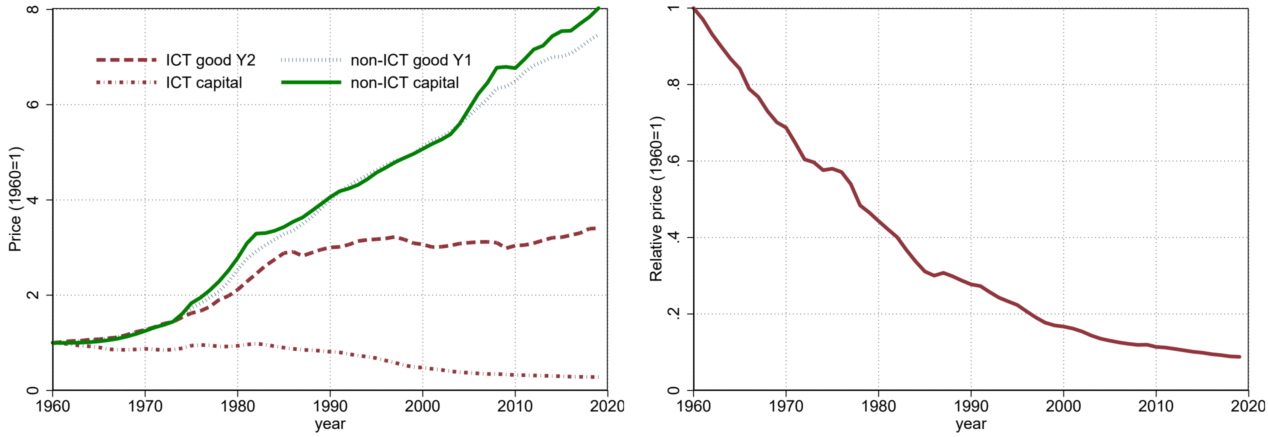
4.1 Advances in Digital Technology

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[21] are classified to be part of the ICT-intensive sector that

[21] 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.

Figure 7: Prices of final consumption goods and capital

(a) Individual price series (b) μ



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 2.1.1. 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.

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 for which we have CEX data, 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 around 17.3%. The non-ICT-intensive sector 1 has an ICT share of 8.3%.

Our model assumptions state that those industries clustered in the ICT-intensive sector are the only producers of ICT capital, whereas they do not produce non-ICT capital. As industries in this sector are producers of ICT goods (such as computer and electronic products, computer systems and design) or service providers (such as as funds, legal services) we conclude that this assumption is a fair reflection of the real world.

According to eq. (10), 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 7 shows the evolution of the prices of Y_1 and Y_2 as well as the prices of ICT capital (I_{IT}) and non-ICT capital (I_K). The series are Törnqvist price indices based on NIPA price series for individual capital classes and commodities.[22] 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 will feed a smoothed version of this series into our simulation as an exogenous change in ICT technology.

4.2 Preference parameters

We estimate the preference parameters θ, ρ and ν directly from our dataset assembled in Section 2. We follow Boppart (2014) by regressing households' logarithmized expenditure share of non-ICT

[22]The NIPA consumer price series are assembled from the more disaggregate BLS data series we use in Section 2.2.

goods on the logarithmized relative price and the logarithmized expenditure level

$$\ln\left(\frac{C_{i,1}P_1}{e_i}\right) = \ln\nu + \theta\ln\left(\frac{P_1}{P_2}\right) - \rho\ln\left(\frac{e_i}{P_2}\right), \tag{23}$$

where he suggests instrumenting expenditure e_i by household income as infrequently bought items may lead to endogeneity in the expenditure level.

We carry out regressions on two different samples. The first sample is the same observational level as in Section 2. We consider 100 income percentiles over 23 years. We prefer to interpret the sample as a repeated cross-section, but we also carry out panel regressions with income percentile fixed effects as a robustness check. Additionally, we estimate eq. (23) at the household level. This allows us to control for household characteristics. As we cannot match households across the diary and interview survey, we restrict our sample to the interview survey. The interview survey covers the whole spectrum of consumer spending, which is why this restriction should be of no consequence. As in section 2.1.3, we compute individual equivalents and truncate the sample to avoid biases due to top-coding in income. We control for a large set of household characteristics: the share of children among household members, the share of old individuals (defined as age 65 and older), whether there are earners in the household other than head and spouse, the sex of the head, the educational level of head and spouse and whether each of them participates in the labor market, and finally whether the household is located in an urban or rural area.

Reassuringly, the regressions at the household and percentile levels result in very similar estimates, as do the OLS and the IV estimates.[23] Our preferred specification is column (5), which yields $\theta = -0.026$, $\rho = 0.025$ and $\nu = 1.094$. Be reminded that ρ controls the strength of the effect of incomes on expenditure shares whereas θ controls the strength of the relative price effect (the effect of technological change). As all estimates are statistically significant, both channels are of empirical importance. ρ is significantly larger than zero, so preferences are non-homothetic and the non-ICT good is a necessity whereas the ICT good is a luxury. This is in line with our finding in Section 2.1.4 that richer households spend a larger share of their budget on ICT goods.

The elasticity of substitution between ICT and non-ICT goods for the average household is $\sigma = 1.524$. 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. ICT and non-ICT goods are partial substitutes. This is in line with our decomposition in Section 2.2 that finds increasing expenditure shares for ICT-intensive products, despite weaker price dynamics.[24]. To our knowledge, our paper is the first to provide estimates of the elasticity of substitution between ICT and non-ICT goods.

4.3 Other parameters

The parameters in the production function cannot be directly estimated from the data, but need to be calibrated. There are four elasticities of substitution (ϵ_i, η_i $\forall i = 1, 2$) and six weights (α_i,γ_i,

[23]When including income fixed effects, we cannot simultaneously use income as an instrument, which is why we do not show an IV fixed effects regression.

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

Table 3: Estimation of preference parameters

	Percentiles			Households	
	(1)	(2)	(3)	(4)	(5)
$\ln(\frac{P_1}{P_2})$	-0.028***	-0.026***	-0.030***	-0.013***	-0.026***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)
$-\ln(\frac{e_i}{P_2})$	0.017***	0.018***	0.015***	0.006***	0.025***
	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)
Constant	0.080***	0.090***	0.066***	-0.085***	0.097***
	(0.009)	(0.010)	(0.014)	(0.004)	(0.010)
Method	OLS	IV	OLS	OLS	IV
Fixed effects	No	No	Yes	No	No
Controls	No	No	No	Yes	Yes
Observations	2,300	2,300	2,300	134,939	134,939
R-squared	0.163	0.162	0.162	0.055	0.025

Note: Results for regression equation (23). Household-level controls include the share of children, the share of age 65+, presence of other earners, sex of the head, education and labor market participation of head and spouse, geographical location. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. (4) and (5) are weighted regressions using the survey weights. The instrument (IV) is income.

ϕ_i), furthermore two depreciation rates δ_i. We make the following simplification: $\epsilon_1 = \epsilon_2 \equiv \epsilon$, $\eta_1 = \eta_2 \equiv \eta$. While this reduces the dimensionality of the calibration exercise, it also intuitively makes sense. The elasticity of substitution reflects the overall flexibility of the economy with respect to using the different inputs, i.e. how easy it is to exchange one input for another. To us, it is straightforward to assume that this is the same in both sectors, since it relates to underlying characteristics of the whole economy. In contrast, the weights determine how much the two sectors rely on each of the inputs in the production process.[25] By allowing γ_i and ϕ_i to differ across sectors, the model is still able to generate sufficient differentiation to match the data series.

We calibrate several parameters outside of the model: The depreciation rates δ_i and capital shares α_i are taken from the BEA. The BEA directly provides estimates for the depreciation of non-ICT and ICT capital. The non-ICT capital share in those industries can be computed by deducting the labor share from overall value added, taking into account the share of non-ICT capital in overall capital in both sectors. The supply of high- and low-skill labor ($\bar{L}$ and $\bar{H}$) is taken from the American Community Survey. We define high-skill workers as workers with college degree and low-skill workers as all others and use the total number of hours worked for each group in 1996.

This leaves us with six parameter values to be calibrated internally. We estimate these using the Simulated Method of Moments (McFadden, 1989). We aim to minimize the sum of squared errors in a set of data moments, scaled by the corresponding moment, which means that we are minimizing the percentage error of moments.

Since we want to pin down the parameters of the production function, we need to choose moments that are informative about the relative use of the inputs L, H, ICT and K (which can move

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

Table 4: Calibration fit

Data moment	Source	Data	Model
Pct. change in skill premium, 1960-2017	ACS	22.3	22.2
Pct. change in relative price, 1960-2017	NIPA	134.6	134.3
Pct. change in ICT intensity, 1960-2017	BEA	2096.2	1719.6
Pct. change in value-added share of sector 2, 1960-2017	BEA	223.7	253.8
Pct. change in labor share, 1960-2017	EHS (2013)	-11.7	-11.5
Pct. change, rate of return to ICT capital, 1960-2017	KN (2019)	-86.0	-90.5

Note: 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.

freely across sectors) and how the two sectors differ in their prices and output produced. The two most important moments are the skill premium –the relative wages of high-skill and low-skill workers from the ACS– and the relative price of goods from Figure 7. They measure the strength of the income and the price channel, respectively. To determine the weights γ_i and ϕ_i, we need to introduce another sector-specific moment. We choose value-added in sector 2 relative to sector 1, which reflects prices and quantities. To match the input factors and their returns, we choose the following additional moments: the ICT intensity – pinning down the use of ICT vs. non-ICT capital –, the return on ICT capital and the labor share, which informs 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. All moments are expressed in percentage changes between 1960 and 2017.

Given our focus on the price vs. income channel, we choose a weighting matrix that puts more weight on the decline in the relative price and the increase in the skill premium. We overweight these moments by a factor of 5. The results are robust to choosing slightly different weights. With six parameters and six data moments, the system is exactly identified under the assumption that the moments are orthogonal.

Table 4 shows the calibration fit. We match most data moments well. The model generates a strong increase in the importance of ICT capital in the economy's total capital stock. This is an important validation of the model as ICT capital was our main empirical measure of digitalization in the previous sections. At the same time, the return to ICT capital drops dramatically, in line with the estimates from the data. The model also matches well the strong growth in value-added in sector 2 relative to sector 1. Despite imposing $\eta_1 = \eta_2$ and $\epsilon_1 = \epsilon_2$, the different weights across sectors make them sufficiently distinct in their use of ICT capital to explain the large differences in value-added growth over time. Finally, the model is able to generate a decline in the labor share of almost the same size as has taken place in the real world.

Table 5: Calibrated parameters

Symbol	Value	Description	Source
Preferences			
γ_1	0.94	Weight high-skill inner nest in 1	estimated by SMM
γ_2	0.25	Weight high-skill inner nest in 2	estimated by SMM
$\epsilon_1 = \epsilon_2$	1.04	El. of Subst between H and ICT	estimated by SMM
Total Work			
ϕ_1	0.08	Weight low-skill middle nest in 1	estimated by SMM
ϕ_2	0.61	Weight low-skill middle nest in 2	estimated by SMM
$\eta_1 = \eta_2$	2.6	El. of Subst betw L and SW in 1	estimated by SMM
Final Production			
$\alpha 1$	0.45	Capital share in 1	BEA
$\alpha 2$	0.18	Capital share in 2	BEA
δ_{ICT}	0.15	Depreciation rate ICT capital	BEA
δ_K	0.065	Depreciation rate non-ICT capital	BEA
Other Parameters			
$\bar{H}$	37.45	Supply of high-skill workers	ACS
$\bar{L}$	76.79	Supply of low-skill workers	ACS
β	0.97	Discount factor	

Table 5 summarizes the calibrated parameters. We see not surprisingly that $\gamma_1 > \gamma_2$ and $\phi_2 > \phi_1$, as sector 2 is more ICT-intensive. As sectors are required to have the same elasticities of substitution, it falls on the weights to generate all the differences across sectors, and in consequence the calibrated values are quite distinct from each other. The large elasticity of substitution between low-skilled labor and SW, η, suggests that low-skilled labor and the ICT capital composite are strong substitutes, whereas ICT capital and high-skilled labor are more complementary to each other due to the lower value of ϵ. As the price of ICT capital falls and the scale of production is increased, more high-skilled labor is used to produce SW, which in turn substitutes for low-skilled labor. This is in line with the existing literature (e.g. Autor and Dorn, 2013, Eden and Gaggl, 2018 and Burstein et al., 2019) and is in particular needed to match the increase in the skill premium that we observe in the data.

5 Results

In this section, we discuss the simulation results in detail, focusing on the effect of digitalization on income and prices of high- and low-skill households. We show that price changes matter by comparing nominal and real expenditures. To estimate the importance of the price effect in terms of welfare, we carry out a counterfactual analysis in form of a compensatory variation, similar to the exercise in the empirical part.

5.1 Income

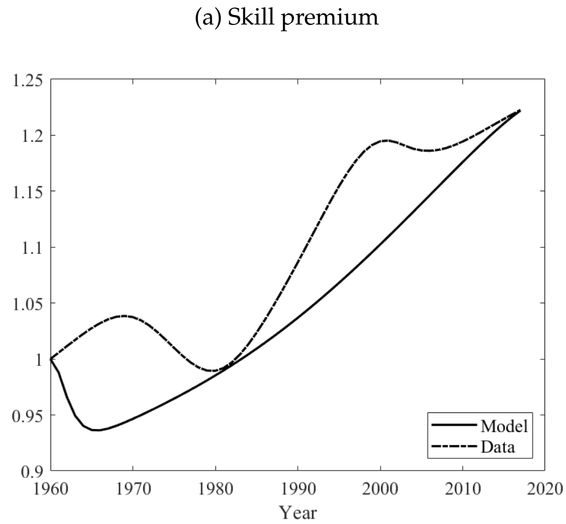
Panel (a) of Figure 8 shows the evolution of the skill premium W_H/W_L, normalized to 1 in 1960. The skill premium increases substantially both in the data and in the model. With the calibrated parameters, we match the long-run increase of the skill premium well. The model does not capture some of the short-run fluctuations during the transition, but this is not surprising given that we focus on long-run trends in the skill demand and abstract from factors that create fluctuations at business-cycle frequencies.

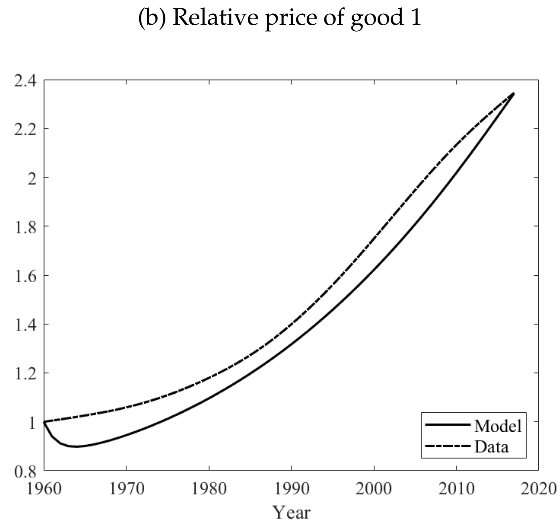
Advances in ICT technology affect the skill premium in the following way: As ICT capital becomes cheaper and more efficient in the production, firms accumulate more ICT capital and increase their ICT intensity. As ICT capital and high-skill labor tend to be complements, demand for high-skill labor rises. Simply put, if a firm uses more ICT capital such as computer hardware or software, it also needs a high-skill worker who can work with that new form of capital. At the same time, low-skill labor can be substituted by the composite input of high-skill labor and ICT technology. Spoken simply again, certain tasks can be executed more efficiently by computers operated by high-skill workers, rather than by low-skill workers.

As we assume the supply of both high- and low-skill labor to be fixed, the changes in the demand for workers will manifest entirely in a change in wages. Let us consider separately the evolution of high- and low-skill wages. We calculate real wages by deflating with an aggregate Törnqvist price index. High-skill workers benefit strongly from digitalization. Their real wage W_H rises by 29.0%. The effect on low-skill wages is a priori ambiguous: On the one hand, some low-skill jobs are taken over by high-skill workers with ICT equipment. On the other hand, the scale of production increases, as digitalization leads to a boost in productivity. As has been discussed in the literature (e.g. Acemoglu and Restrepo, 2018), the net effect can be either positive or negative depending on the features of the economy and the nature of technological progress. In our case, the real low-skill wage increases by 5.6%, so that the productivity effect dominates. Both skill types are better off due to automation, but high-skill households benefit more.

Wage income is the only source of income for low-skill households, which they fully spend on consumption. In contrast, high-skill households also own the capital stock. So on the one hand, they have an additional source of income in the form of returns to capital, but on the other hand they also need to finance investment. The relevant concept when studying their expenditure patterns is therefore disposable income, i.e. wage plus capital income minus investments. Disposable income of high-skill households increases by 28.2% over the simulation period. This number is almost identical to the increase in the wage because of two counter-acting forces: On the one hand, capital income grows by more than high-skill wage income. On the other hand, the share of investment in expenditures of high-skill households increases. As a result, the share of wage income in net income stays almost constant over time. Given these numbers, the increase in income inequality between high- and low-skill households in our model is almost completely driven by the increase in the skill premium and not by larger capital income.

Figure 8: Simulation results





Note: The figures plot the model (solid line) against the data, where available (dotted line). Data and model series are normalized to 1 in 1960. Panel (a) shows the evolution of the skill premium W_H/W_L. Panel (b) plots the relative price as the price of the non-ICT-intensive good 1 relative to the ICT-intensive good 2. *Data sources:* ACS and NIPA.

5.2 Relative prices and relative expenditures

Panel (b) of Figure 8 shows how the relative price P_1/P_2 increases in response to the technology shock. Although not targeted, the model captures the transition dynamics between 1960 and 2017 well.

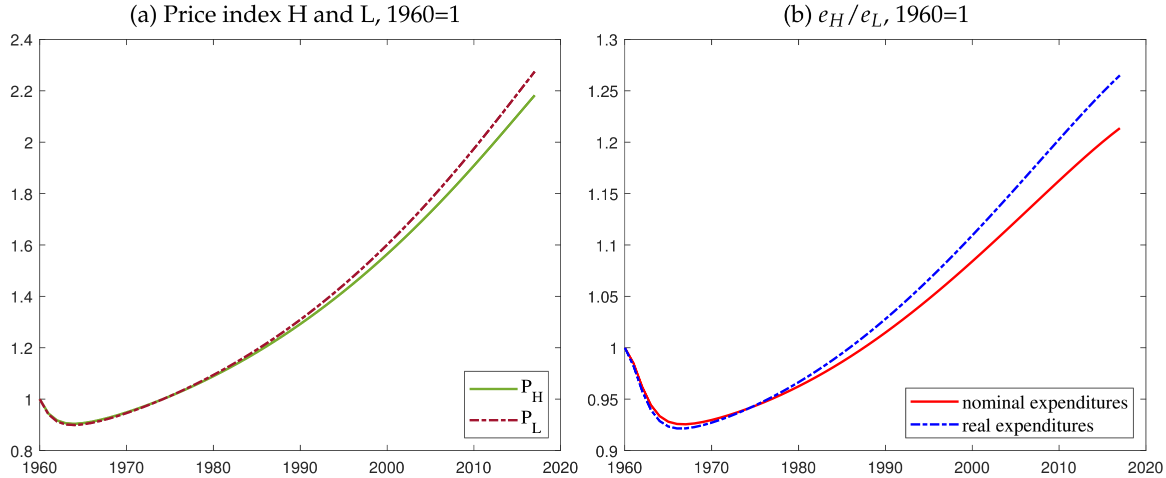
Progress in digitalization favors sector 2, as it relies more on ICT capital as a factor input. While the declining price of ICT capital implies that it gets cheaper to produce both goods, sector 2 experiences larger productivity increases. In consequence, the relative price of good 1 rises. Due to higher demand for ICT capital, investment activity shifts from non-ICT capital to ICT-capital.

Consumption demand also shifts towards good 2. Expenditure for good 2 in total consumption expenditures almost doubles between 1960 and 2017. This is not solely due to the change in the relative price, but it is also linked to non-homothetic preferences. As households get richer, they want to consume a larger share of good 2 even in absence of any price changes. In principle, lower relative demand for good 1 puts downward pressure on its relative price. In our model the effect of technological progress is more important and therefore the relative price of good 1 still increases substantially.

To assess the strength of the price effect, we consider changes in the overall cost of living. A comprehensive consumer price index takes changes in both P_1 and P_2 into account and differs across households because high- and low-skill households consume different bundles of goods. We therefore calculate separate Törnqvist price indices for each skill type using their specific expenditure shares of good 1 and 2.

Figure 9 (a) shows the evolution of the household-specific consumer price index. The costs of living increase for both households, but importantly, the index for low-skill households increases by a larger margin (2.5% more) as their consumption bundle depends more on non-ICT goods.

Figure 9: Price indexes and expenditure of high- vs. low-skill households



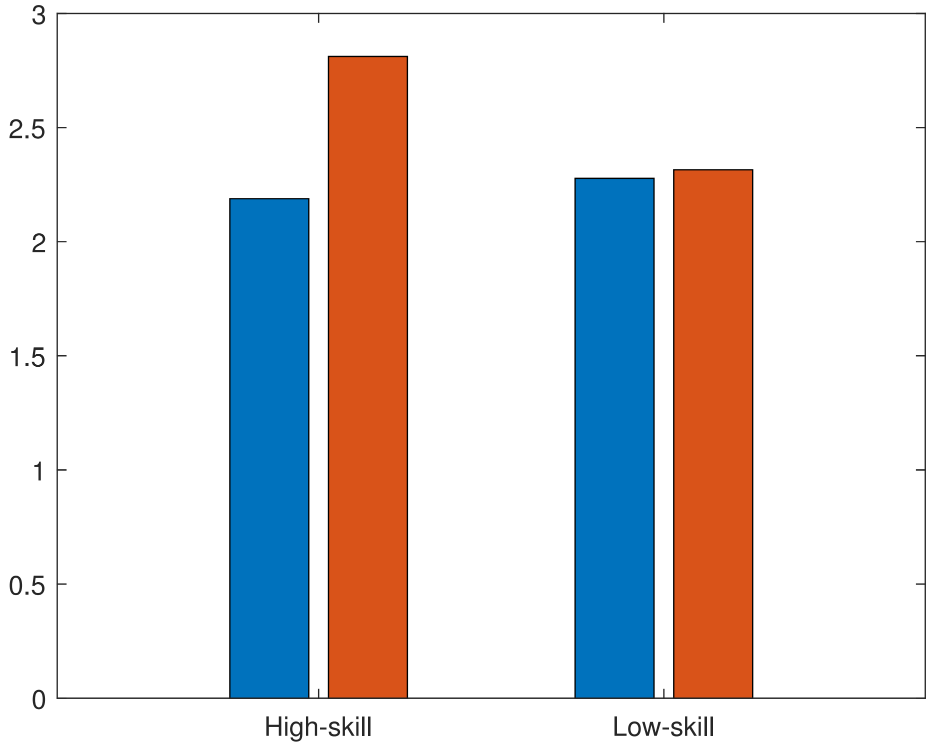
Note: Panel (a) shows how both personal Törnqvists price indices evolved over time. Panel (b) shows the ratio of nominal expenditures (red solid) and the ratio of real expenditures $e_H/P_H/(e_L/P_L)$ (dashed blue) in the model.

Figure 9 (b) tracks the evolution of expenditure of high-skill relative to low-skill households, and is a measure of consumption inequality. The solid red line shows the ratio of nominal expenditures. Nominal expenditures of high-skill households increase by 22% more than those of low-skill households, in line with the increased skill premium of 22% in Figure 8. The dashed blue line shows the ratio of real expenditures for consumption, defined as nominal expenditures divided by the household-specific price index. When taking into account the changed purchasing power of households, real consumption expenditures of high-skill households increase by 26%, a plus of four percentage points compared to the pure nominal measure. The relative price changes therefore augment the difference in consumption inequality by around a quarter, which means that the price channel is an important driver of consumption inequality. This difference is in line with the back-of-the-envelope calculation using the numbers presented in the empirical part of our paper: In the data analysis the increase is 2.8% for the considered time horizon, broadly in line with the divergence of the price index in the model of 2.5%.

Our model results are in line with several recent papers that have documented an increase in consumption inequality in the U.S. over the last decades[26]. The discussion often centers on whether consumption inequality has mirrored income inequality in nominal terms. For example, Aguiar and Bils (2015) find that consumption inequality – measured as nominal consumption expenditures – has increased by around 30% between 1980 and 2010 – in line with increases in income inequality. We add to the literature by establishing digitalization as one driving force of long-run dynamics in consumption inequality, working through increases in income inequality and changes in the relative price of different goods. We emphasize that real consumption inequality increases by an even larger margin. Focusing narrowly on nominal consumption expenditures understates the overall effect.

[26]E.g. Aguiar and Bils (2015) and Attanasio and Pistaferri (2014); see Attanasio and Pistaferri (2016) for an overview.

Figure 10: Compensatory Income vs. actual income



Note: The figure shows how much income households get with the compensatory variation (blue) and how much income households actually get in 2017 (red) relative to the initial 1960 income.

5.3 Welfare effects of digitalization

In the previous section, we have established that both the income and the price channel favor high-skill households, and we have measured the relative size of each channel. This section aims to assess the welfare effects of digitalization for high- and low-skill households via a counterfactual exercise.

We conduct a compensatory variation, which asks how much additional income households need 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. We thus compute in monetary units how much more high-skill households benefit from digitalization than low-skill households. This procedure is analogous to our empirical analysis in Section 2.2, but avoids the shortcomings of the pure data analysis.

For the compensatory variation, we consider the initial welfare level of households as of 1960. This equals the utility u_j from eq. (12), using the actual expenditures in 1960. Then, we assume that prices in 1960 were as in 2017. The compensatory variation indicates the expenditures needed to achieve the same level of utility as in 1960, u_j, using the 2017 prices. In the next step, we compare this compensatory measure with the actual increase in income that households experience between 1960 to 2017 in the model. If the actual increase in income is larger than the compensation, households are better off in the post-digitalization world. We do this exercise separately for high-skill and low-skill households.

Figure 10 compares the compensatory income with the actual income for high- and low-skill households. The initial income is normalized to 1 for each skill type. This means that the numbers in the graph are to be interpreted relative to 1960 nominal income. The blue bar shows the compensatory income. As low-skill households rely more on consumption of good 1 than high-skill households, their compensatory income is larger (227.8% of income in 1960, compared to 218.8% for high-skill households). The red bar shows how much more nominal income households actually receive in 2017 relative to 1960. As digitalization increases wages of high-skill households more, their disposable income increases by 281.2%. In contrast, low-skill household's nominal income increases by 231.5%.

This means that high-skill households experience a substantial increase in welfare due to digitalization, whereas low-skill households gain almost nothing. In terms of additional income, high-skill households gain welfare equivalent to 28.5% of their initial income (281.2/218.8 -1), whereas low-skill households gain only 2%. This large difference is because of an increase in wage polarization, but also due to the price response of goods that disproportionately favors the rich. If both households had the same consumption behavior, the required compensation for price changes would have been the same, relative to their initial income. This makes the computed welfare gains more equal. For example, with the same compensatory income for high-skill households as for low-skill households their welfare gain would only be equivalent to 23.4% instead of 28.5%. This is 18% less and is broadly in line with the 26% difference in real consumption expenditures when considering price changes in the model, see the previous subsection.

This brings us back to the motivation of this paper. The literature on the effect of digitalization on income inequality has provided important insights and taking income changes into account also goes a big step towards assessing the welfare effects of digitalization. But to capture the full welfare effects, price changes need to be taken into account as well.

6 Conclusion

Households differ in the digital share of their consumption basket. This paper proposes a measure of digitalization in consumption, which relies on the share of digital assets in the capital stock of the production process. Combining this measure with consumption data along the income distribution, we show that rich households consume a larger share of digitally-produced goods than poor households. Furthermore we provide evidence that the price change of ICT-intensive relative to non-intensive commodities has been more beneficial for rich households. We present a structural model, in which digitalization affects consumption both via a change in incomes and via a change in the prices for consumption goods. Calibrating the model to the US economy between 1960 and 2017, we quantify the effect of the substantial decline in ICT capital prices on consumption inequality. We find that inequality of real consumption has increased by 26%. Around a quarter of this increase is attributed to changes in relative prices. These 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.

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.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy 128*(6), 2188–2244.

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.

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

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., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics 118*(4), 1279–1333.

Barkai, S. (2020). Declining Labor and Capital Shares. *Journal of Finance 75*(5), 2421–2463.

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.

Fajgelbaum, P. D. and A. K. Khandelwal (2016). Measuring the unequal gains from trade. *Quarterly Journal of Economics 131*(3), 1113–1180.

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. *The 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.

Hochmuth, P., M. Pettersson, and W. Christoffer (2022). A Nonhomothetic Price Index and Inflation Heterogeneity. *Unpublished Working Paper*.

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. (2018). The race between preferences and technology. *Unpublished Working Paper*.

Jaimovich, N., I. Saporta-Eksten, H. E. Siu, and Y. Yedid-Levi (2020). The macroeconomics of automation: Data, theory, and policy analysis. *NBER Working Paper No. 27122*.

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

Jaravel, X. (2021). Inflation inequality: Measurement, causes, and policy implications. *Annual Review of Economics 13*, 599–629.

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 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.

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 (2021). Benign effects of automation: New evidence from patent texts. *Review of Economics and Statistics* (forthcoming).

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*, 995–1026.

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 (2021). Uneven growth: Automation's impact on income and wealth inequality. *NBER Working Paper No.21595*.

Muellbauer, J. (1975). Aggregation, income distribution and consumer demand. *Review of Economic Studies 42*(4), 525–543.

Nigai, S. (2016). On measuring the welfare gains from trade under consumer heterogeneity. *Economic Journal 126*(593), 1193–1237.

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.

A Data Appendix

A.1 Details on data sources and data construction

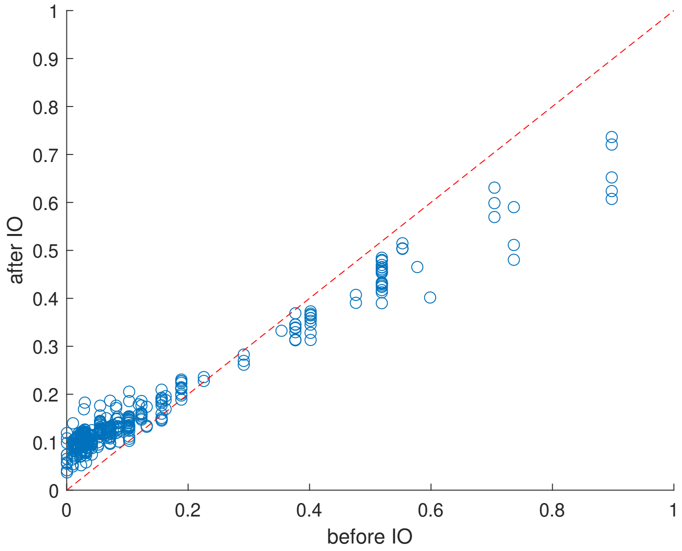
Input-Output Accounts

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 sum of the digitalization shares of its intermediate inputs and of value added.

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.

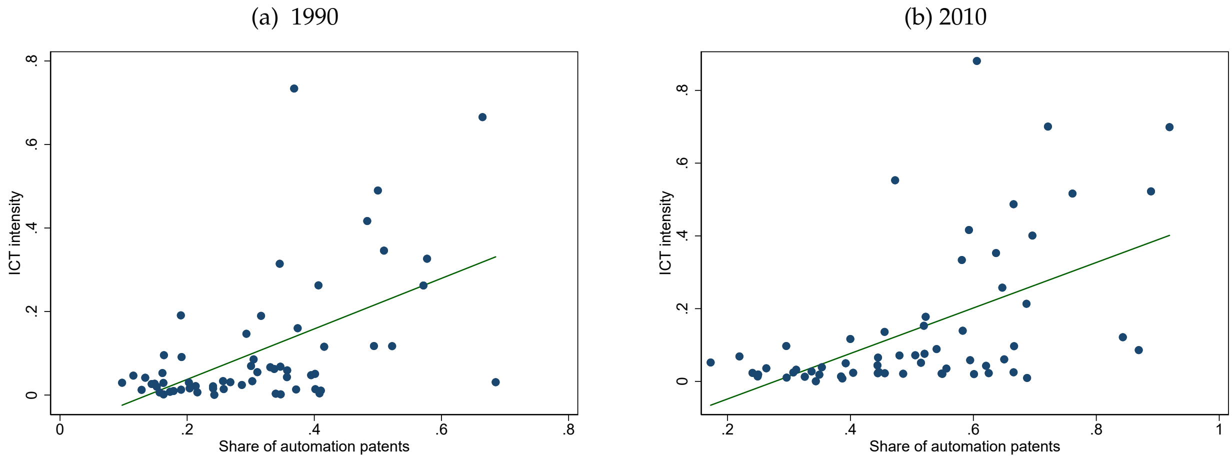
One dimension of inequality that we cannot capture with our data are quality differences in the goods consumed. A rare exception is the sub-category restaurants (broad industry 722, corresponding to "food away from home" in CEX). Households in the CEX are asked to report separately money spent at fast 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.

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 (2021). Mann and Püttmann (2021) 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[27]. 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.

[27] The plots would look similar for other years.

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

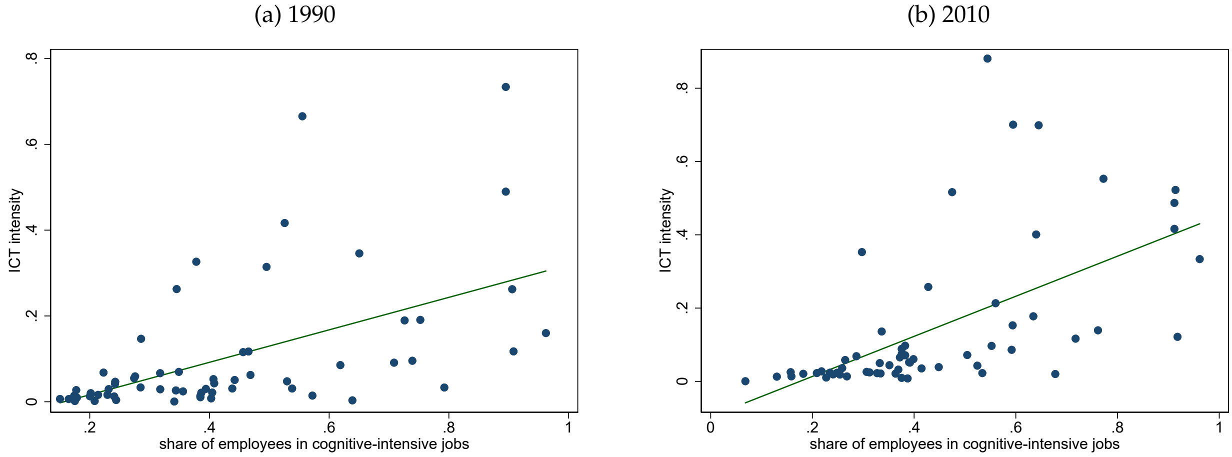


Note: The graph plots ICT intensity against the share of automation patents in the total number of patents by BEA industry. Each dot represents one industry, the line shows a linear prediction. The left panel shows raw data for 1990, the right panel for 2010. *Source:* Mann and Püttmann (2021), BEA and own calculations.

Another established procedure in the literature is to consider the input of different tasks across occupations in each industry. Autor et al. (2003) and Autor and Dorn (2013) focus on routine tasks as a measure of automation potential of production processes. Gaggl and Wright (2017) and Adão et al. (2020), among others, argue that in the context of ICT, it is more relevant to focus on cognitive tasks, which are presumably complementary to ICT capital. We define a cognitive-task intensity analogous to the routine-task intensity of Autor and Dorn (2013), by taking the log of abstract tasks divided by routine and manual tasks by occupation and defining a job as cognitive-intensive when this number is in the upper third of the distribution.

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

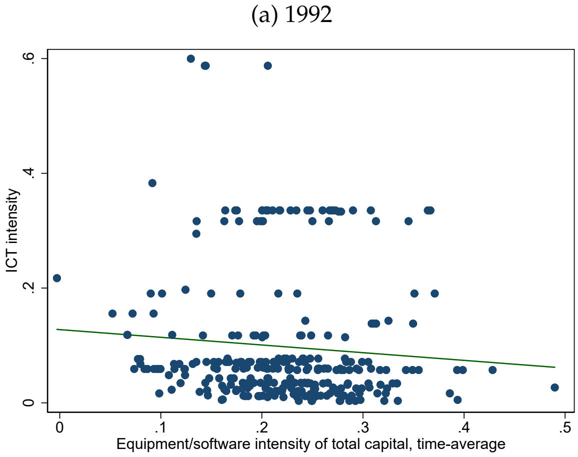
Figure A3: Relationship between share of cognitive employment and the share of ICT capital on overall capital

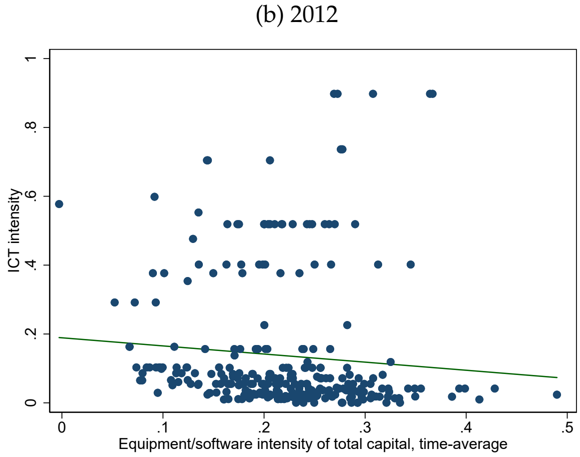


Note: The graph plots ICT intensity against the share of employees that work in cognitive-intensive jobs by BEA industry for 1990 (left panel) and 2010 (right panel). Each dot represents one industry, the line shows a linear prediction. *Source:* American Community Survey, Autor and Dorn (2013), BEA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

Some papers use a similar approach to ours by considering the capital stock at the industry level. Hubmer (2018) 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 (2018), 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 (2018) 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 A4: Relationship between the share of overall equipment and software on capital as in Hubmer (2018) and the share of ICT capital on overall capital





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 (2018), BEA and own calculations.

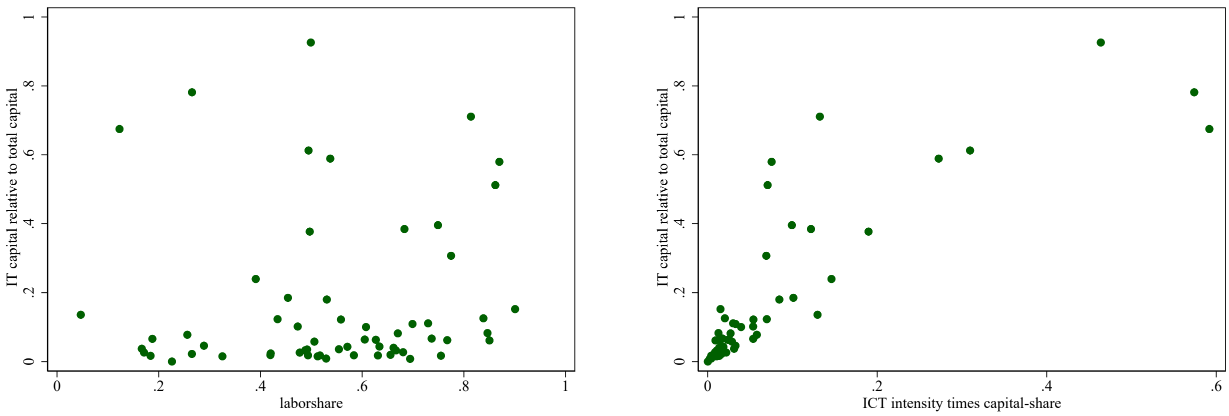
A large body of literature focuses on robots when measuring automation technology (see e.g. Graetz and Michaels, 2018, Acemoglu and Restrepo, 2020). We have also computed a robot intensity using data from the International Federation of Robotics and correlated it with our ICT intensity. As robots are not widespread and are only used in a few industries, which have little overlap with ICT-intensive industries, the correlation to our measure is very low (results available upon request).

Lastly, we assess in how far the labor share is related to our capital intensity measure. The following figures illustrate the correlation between the labor share and ICT intensity in 2017 and our alternative measure defined as ICT intensity times one minus the labor share. We conclude that the labor share is not systematically related to our measure on the level of the BEA industries. We therefore proceed with our original measure of ICT intensity.

Figure A5: Relationship ICT intensity to labor share

(a) ICT intensity and the labor share

(b) ICT intensity and labor share adjusted measure



Note: The graph plots ICT intensity against the labor share (left panel) and the ICT intensity times one minus the labor share (right panel). Each dot represents one industry on the BEA level for 2017. *Source:* BEA and own calculations.

A.3 Input- vs. output-based digitalization measures

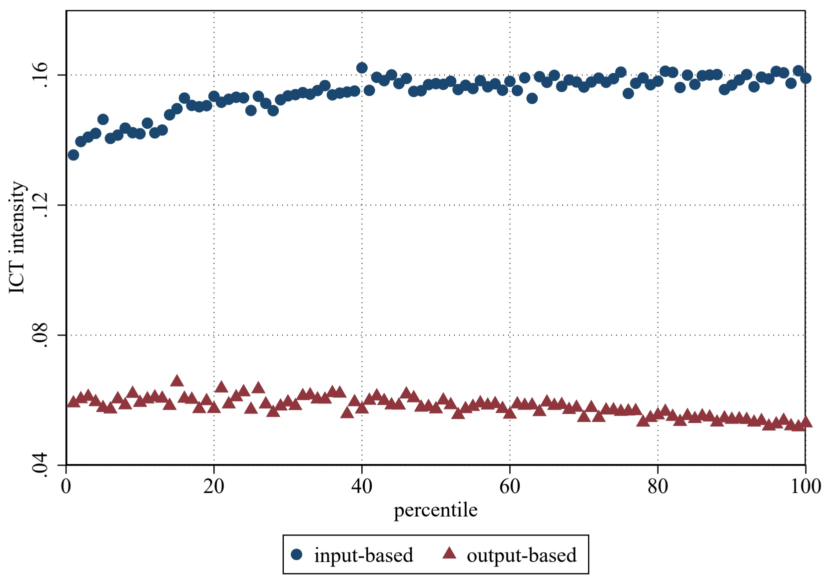
In the main text, we define digitalization based on the inputs used in the production process. A good is classified as digital according to the extent to which it has been produced with digital assets like computer hardware or software. An alternative would be an output-based measure, sorting final consumption 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 A6 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. Figure A7 plots the output-based ICT intensity against consumer price inflation. As in Figure 5, the correlation is negative, which implies that households with a more ICT-intensive basket also fare better when focusing on output-based digitalization. Combining the findings of these two graphs, we would conclude that poorer households tend to be the beneficiaries from digitalization when measured based on outputs.

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.

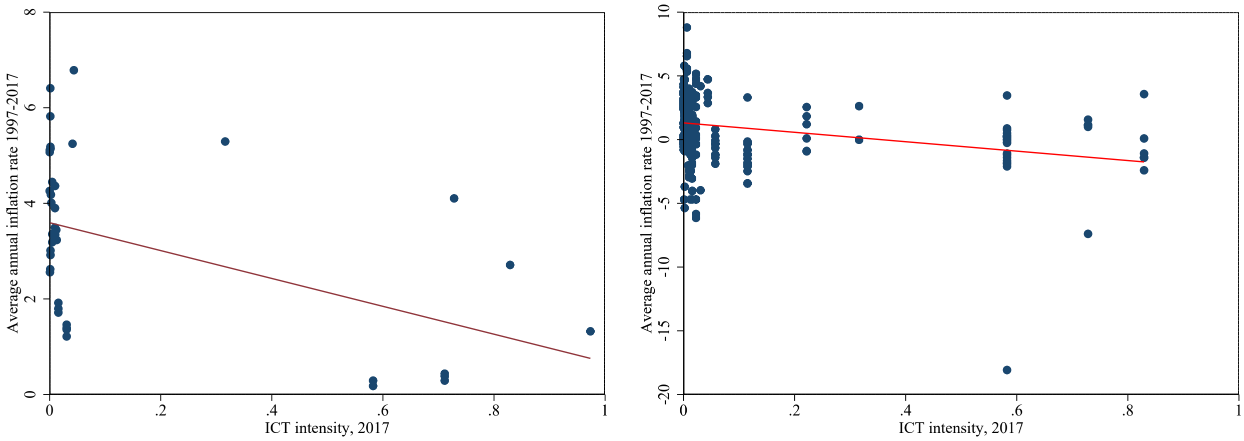
Figure A6: ICT intensity along the income distribution, output-based classification



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.

Figure A7: Consumer price inflation and ICT intensity, output-based classification

(a) BEA industries (b) IO commodities



Note: The graph plots the ICT intensity in 2017 against average annual inflation in consumer prices between 1997 and 2017. Each dot represents a consumption category (left: BEA industries, right: IO commodities). The red line shows the prediction from a linear regression.*Source:* BEA, CEX, BLS and own calculations.

A.4 Inflation rate and ICT intensity, robustness

Table A1 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. In column (2), we repeat the regression but cut the sample as 2012 in order to make the results comparable to 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.

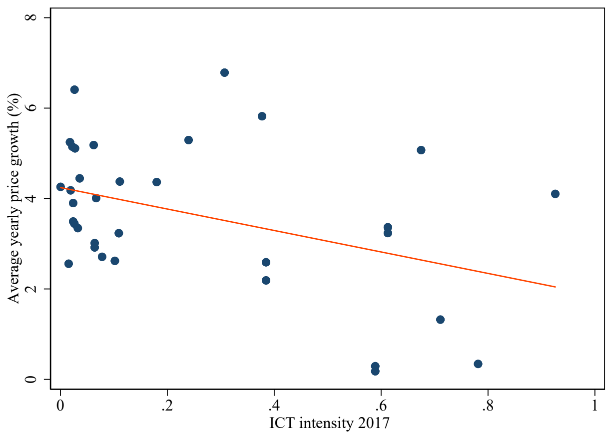
Table A1: Average annual inflation rate and ICT intensity

Variables	(1) $\bar{\pi}_{2017}$	(2) $\bar{\pi}_{2012}$	(3) $\bar{\pi}_{2012}$
ICT intensity	-2.424***	-4.305***	-2.612***
	(0.792)	(0.944)	(0.871)
import share			-2.933***
			(0.273)
Constant	1.488***	1.725***	2.265***
	(0.161)	(0.170)	(0.162)
Observations	552	536	536
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.

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.

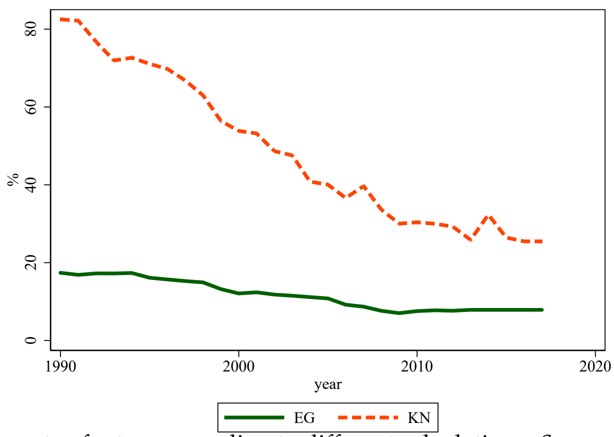
Table A2: Inflation decomposition for 1st and 10th decile

year	π_t		π_t^{fixed}		Δ	
	1st	10th	1st	10th	1st	10th
1996	-	-	-	-	-	-
1997	0.9	1	0.9	1	0	0
1998	1	1.2	1	1.2	0	0
1999	2.6	2.6	2.4	2.5	0.2	0.1
2000	3.1	3	2.9	2.8	0.2	0.2
2001	0.9	1	0.9	1	0	0
2002	2.2	2.3	1.8	2	0.4	0.3
2003	1.3	1.4	1.4	1.3	-0.1	0.1
2004	3.4	3.1	2.9	2.8	0.5	0.3
2005	3.3	2.9	2.9	2.8	0.4	0.1
2006	2.2	2.4	2	2.2	0.2	0.2
2007	4.6	4	3.8	3.4	0.8	0.6
2008	-1.4	-0.5	0.9	0.4	-2.3	-0.9
2009	3.6	3.2	2.1	2.6	1.5	0.6
2010	1.7	1.5	1.4	1.3	0.3	0.2
2011	3.5	3.1	3.3	2.9	0.2	0.2
2012	1.6	1.8	1.3	1.6	0.3	0.2
2013	1.3	1.4	1.1	1.3	0.2	0.1
2014	0.1	0.4	0.9	0.7	-0.8	-0.3
2015	-0.3	0.4	-0.4	0.4	0.1	0
2016	1.3	1.8	0.7	1.4	0.6	0.4
2017	1.9	1.7	1.5	1.6	0.4	0.1

Note: The table shows the decomposition of inflation for the first and the tenth income decile. Inflation π_t is the inflation rate of the price index with changing weights and π_t^{fixed} is inflation with fixed 1996/1997 weights. Rounding errors explain residuals in the decomposition. *Source:* BEA, CEX and own calculations.

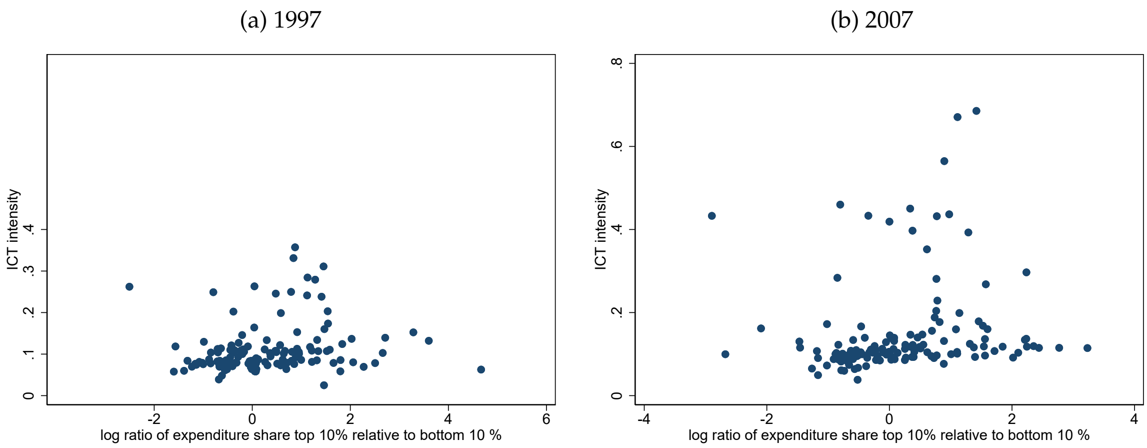
A.5 Additional figures

Figure A9: Rate of returns of ICT capital according to different measures



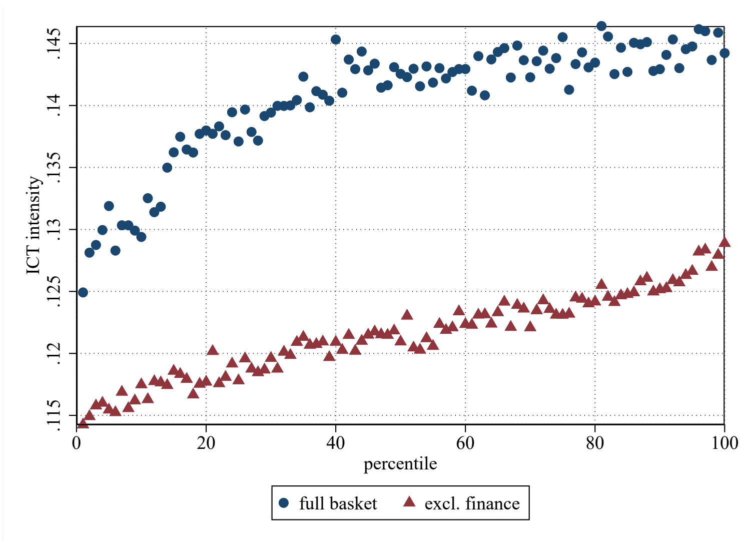
Note: The graph plots the rate of return according to different calculations. *Source:* KN=Karabarbounis and Neiman (2014) and EG=Eden and Gaggl (2018)

Figure A10: ICT intensity of consumption, top 10% relative to bottom 90%



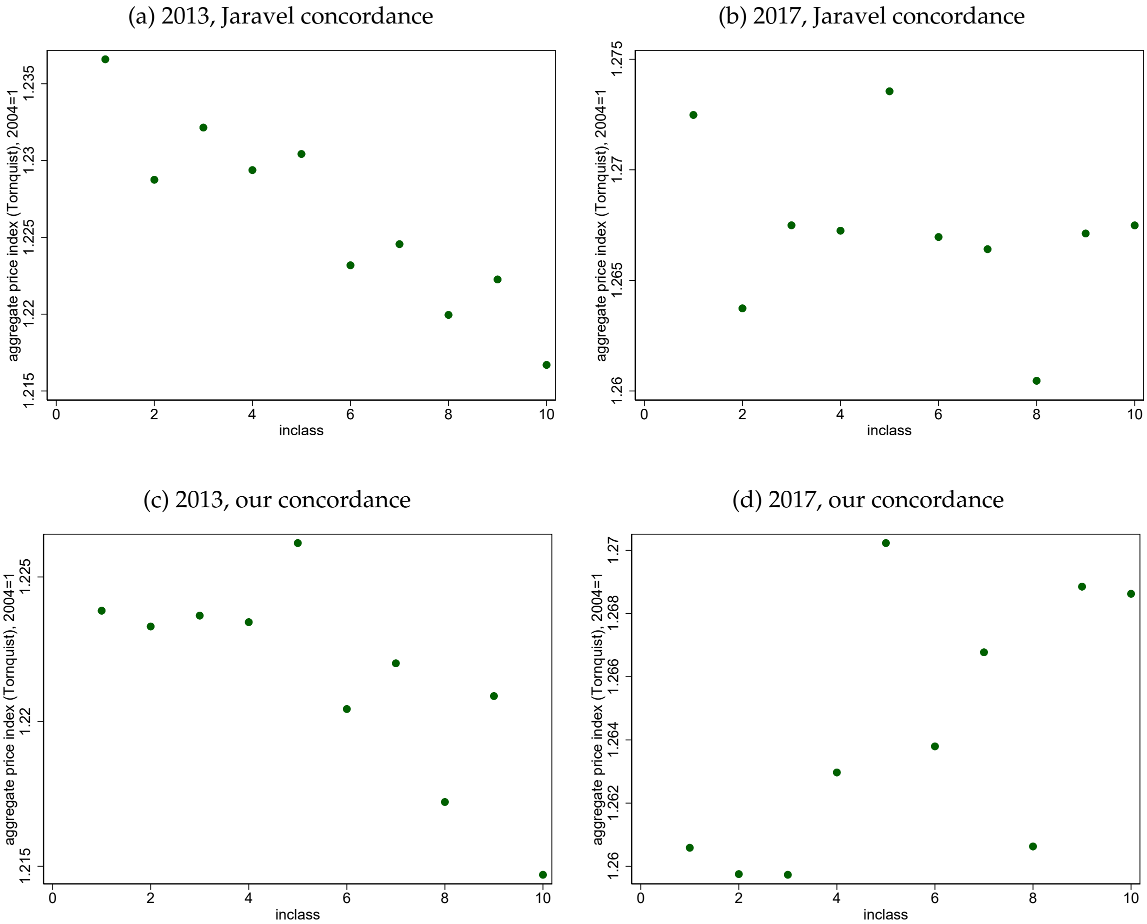
Note: The graphs plot ICT intensity against the log ratio of expenditures of the top 10 versus the bottom 10 percent for each category in 1997 and 2017. *Source:* BEA, CEX and own calculations.

Figure A11: 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.

Figure A12: Consumer price changes across income deciles



Note: The graph shows the price index for each income decile between 2004 and 2013 and 2017 respectively. The first row shows the price index with the CEX-BLS concordance of Jaravel (2019), while the second uses our concordance. *Source:* BLS, CEX and own calculations.

A.6 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.7 Income and asset data from the SCF

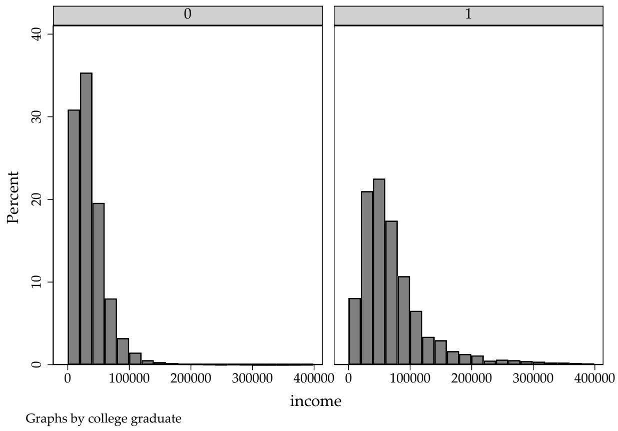
The aim of this section is to motivate our assumption that capital is owned only by high-skill households. To show that this assumption is supported by data, we use data from the Survey of Consumer Finances (SCF) between 1998 and 2013. We focus on households where the head is between 20 and 64 years 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 A3: 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

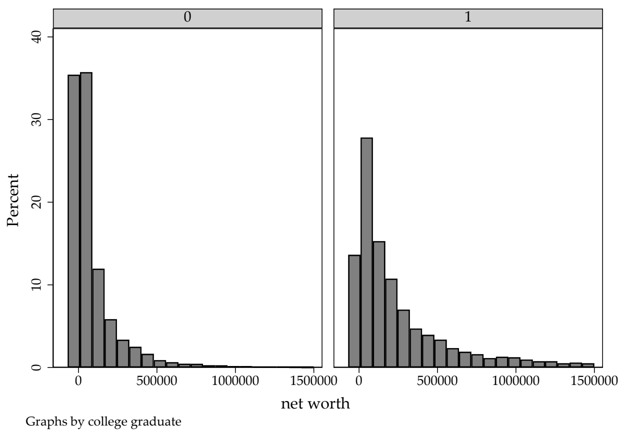
All numbers are in 2019 US Dollars and refer to the median household in each education group. Total income is the sum of labor income, capital income and transfer income.

Figure A13: 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 A14: 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.