

Arvai, Kai; Mann, Katja

Conference Paper

Consumption Inequality in the Digital Age

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Arvai, Kai; Mann, Katja (2022) : Consumption Inequality in the Digital Age, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/264001>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Consumption Inequality in the Digital Age*

Kai Arvai[†]

Banque de France

Katja Mann[‡]

Copenhagen Business School

23 December 2021

Abstract

This paper studies the effect of digitalization on consumption inequality. We assemble a novel dataset of digital technology in the consumption basket of US households and establish a new stylized fact: High-income households have a higher consumption share of digital 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 digital 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 explains 22.5% of the increase in consumption inequality.

Keywords: Digitalization, Inequality, Consumption, Income

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

*We would like to thank Keith Kuester and Moritz Kuhn for helpful comments and suggestions. We are also grateful to Xavier Jaravel and Kirill Borusyak for sharing their CEX concordance with us. The views in this paper are the one 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 is transforming our economy, as it fundamentally changes the way we consume and produce. Digitalization and automation more broadly have been associated with increasing income inequality in the Western world. The literature documents that the increased usage of automation capital, such as robots and computers, has contributed to wage and employment polarization (Acemoglu and Autor, 2011; Autor and Dorn, 2013; Hémous and Olsen, 2020) and a fall in the labor share (Eden and Gaggl, 2018; Acemoglu and Restrepo, 2018; Martinez, 2018), even though it may have been beneficial for the overall number of jobs (Mann and Püttmann, 2021; Gaggl and Wright, 2017). While the debate over rising levels of inequality centers around income, what ultimately matters for welfare is consumption. A comprehensive assessment of the welfare consequences therefore has to consider not just the income effect of digitalization, but also its effect on prices of different types of goods.

A priori, it is unclear how this price effect should impact consumption inequality. If 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. This effect may then either work in the same direction or in the opposite direction of the income effect. The current paper establishes that high-income households have a higher digital share in their consumption basket and are the main beneficiaries from declining prices of digital products. Thus, the price channel amplifies the effect of digitalization on inequality. We obtain this result in a rigorous empirical analysis using household- and sector-level data. We then go on to quantify the importance of each channel in a structural model calibrated to the US economy between 1960 and 2017. While the income effect is larger, the price channel has sizable welfare effects and explains 22.5% of the increase in consumption inequality that occurs due to digitalization.

The paper consists of two parts. In the first part, we construct a new measure of the digitization 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 measure their share in industry-level capital stocks. The resulting digitalization measure covers 61 industries between 1960 and 2017. We find that the share of ICT capital in the overall capital stock has increased from almost 0% in 1960 to over 16% in 2017. There is substantial heterogeneity in the usage of ICT capital across industries. We account for the digitalization content of intermediate products by relying on the input-output structure of the production network. As a next step, we link this industry-level digitalization measure of final goods 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. In particular, consumption categories such as food or textiles, which are consumed disproportionately by poor households, are produced by industries that have a low ICT share. Categories that tend to be more important for rich households, such as finance and insurance or education, have higher ICT shares. We also document that consumer price inflation is weaker for ICT-intensive commodities, which means that digitalization benefits consumers of ICT-intensive goods.

In the second part of the paper, we construct a structural model building on these findings. The model features a two-sector economy with two types of capital, ICT capital and non-ICT capital. Sector 2 uses ICT capital more intensively than sector 1. The economy is populated by two types of agents that differ by skill endowment. High-skill labor is complementary to ICT capital, whereas a composite good constructed from these two inputs is substitutable to low-skill labor. Digitalization is modeled as an increase in the rate of transformation of output into ICT capital. The ICT-intensive sector benefits relatively more from this technology trend, which means that the relative price decreases. At the same time, the skill premium increases. In a setting with non-homothetic preferences, the effect of changing relative prices depends on the agent's position in the income distribution. In line with our empirical finding, we assume that the ICT-intensive good is the luxury good, which is consumed more intensely by the high-skill, high-earner households. To assess the relative importance of income and price changes for consumption, we calibrate the model to the U.S. economy between 1960 and 2017. The simulated increase in income inequality goes along with a decline in prices for ICT intensive goods, benefiting the high earners. In an analysis of the overall welfare effects, we show that high-skill households experience an increase in welfare that is equivalent to 22.3% of their initial income, while low-skill households only experience an increase of around 5.3%. This large difference is due to increased wage polarization and due to the price effect, both favoring the rich. A counterfactual analysis reveals that if relative prices have stayed the same as in 1960, consumption inequality would have increased by 22.5% less than with the price change.

Related literature: This paper is related to the literature on automation and wage polarization. Most of these papers define automation more broadly than digitalization, also comprising automated machines like robots. Autor et al. (2008), Acemoglu and Autor (2011), Autor and Dorn (2013), Hémous and Olsen (2020) and Burstein et al. (2019), among many others, show that recent technological change goes to the detriment of low-skill or routine-intensive occupations. Computers and robots substitute for low-skill workers and complement high-skill workers. As automation technology becomes more productive, this leads to an increase in the skill-premium. Cortes et al. (2017) document a decline in the propensity to work in routine jobs as well. In the life-cycle model of Prettnner and Strulik (2020), automation not only leads to higher income inequality, but also to a more unequal distribution of wealth. Moll et al. (2021) also show that automation increases wealth inequality, in their setting working through higher returns to capital. Related to this finding, Eden and Gaggl (2018), Acemoglu and Restrepo (2018) and Martinez (2018) argue that automation depresses the labor share in value added. The production side of our model is most closely related to Eden and Gaggl (2018), who build on Krusell et al. (2000) to show that automation exacerbates income inequality by increasing the skill-premium and the income share of capital. We rely on the literature on sector-biased technological change (as summarized recently by Herrendorf et al., 2014) to extend the framework to a two-sector economy where technological progress has unequal effects across sectors. We then continue to emphasize the effect of digitalization on goods prices and its impact on consumption inequality.

In studying the effect of digitalization on consumption, we also relate to the literature on consumption inequality. Aguiar and Bils (2015) and Attanasio and Pistaferri (2016) document that consumption inequality closely tracks income inequality when correcting for measurement error

in survey data. Our paper also points out that consumption inequality reflects changes in incomes, but we also emphasize that price effects matter. Hubmer (2018) 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. He links this finding to the declining labor share. While consumption inequality has not yet been in the focus of the automation literature, it has been studied by the trade literature, e.g. Fajgelbaum and Khandelwal (2016), Nigai (2016) and Borusyak and Jaravel (2018). The effect of automation is similar and different to 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. Our paper adopts the approach taken by the trade literature and in particular Borusyak and Jaravel (2018) and is the first – to our knowledge – to study the distributional consequences of automation technology in consumption.

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 relative price channel. We assemble a novel dataset and establish that households along the income distribution differ in the digitalization share of their consumption basket. Our data show that the rich consume more ICT-intensive goods than the poor. At the same time we find that prices of ICT goods have grown at a slower pace than prices of non-ICT goods, making rich households beneficiaries of digitalization.

We assemble our dataset by focusing on industry-level data on the capital stock and household-level consumption data and proceed along the following steps: We (1) create an industry-level measure of ICT intensity by computing the ICT vs. non-ICT capital used in production. (2) trace linkages across industries to create an ICT intensity at the level of final commodities. (3) match commodities to consumption good categories and calculate the digitalization share of consumption baskets along the income distribution. (4) using supplementary data on prices, document a correlation between ICT intensity and consumer prices.

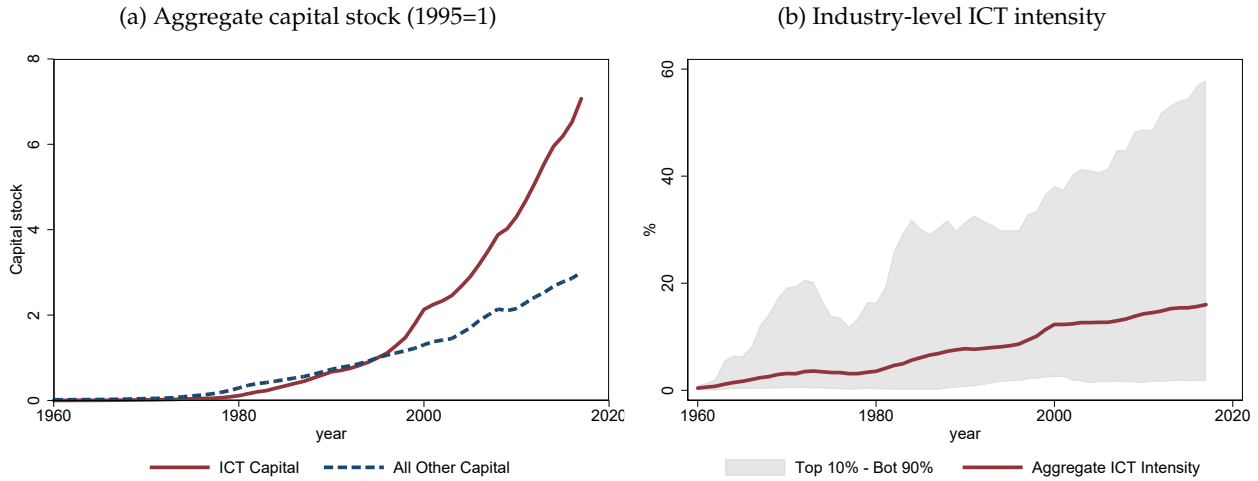
2.1 ICT capital share by industry

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. 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. The same approach is chosen by Eden and Gaggl (2018) and Aghion et al. (2020).

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



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.

Throughout the paper, we will 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**. In adopting this definition, we focus on (relative) inputs in the production process, rather than, for example, the income shares of different factors of production. We argue that the ICT intensity best reflects the structure of the production process, thereby capturing the role that digital capital play in producing certain goods and services. Since both types of capital are evaluated at their current prices, the measure describes how much the ICT capital is worth to producers relative to the non-ICT capital. The relative valuation reflects how productive each capital type is, i.e. the level of ICT- and non-ICT technology. An additional advantage of our digitalization measure is that all data series are directly observable.¹

Ignoring any interlinkages between industries, the right panel of Figure 1 shows the ICT intensity by industry and focuses on the aggregate ICT intensity as well as industries in the 10th and 90th percentile. 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. Industries in the top 10% of the ICT

¹This is not the case when focusing on the ICT capital income share in value added, which has been used in some papers to measure digitalization, see e.g. Eden and Gaggl (2018) and Karabarbounis and Neiman (2019). These measures rely on estimated rates of return to ICT capital, which differ widely across papers.

intensity mostly belong to the finance and insurance industry or to the computer and electronic products manufacturing industry. The least ICT-intensive industries are for example agriculture, the plastic and rubber industry or 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.

2.2 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 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 redefinitions and focus on private industries. The tables contain 394 private industries in 2012 and 473 private industries in 1997, which can be matched to the 61 BEA industries². 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 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 I-O commodities. We focus on this time period, because later on we will link the industry data to the CEX, which is available only from 1996. Commodities that initially have a low degree of digitalization tend to be more digitized after the inputs from other industries are considered. The reverse is true for highly digitized industries (see Figure A1).

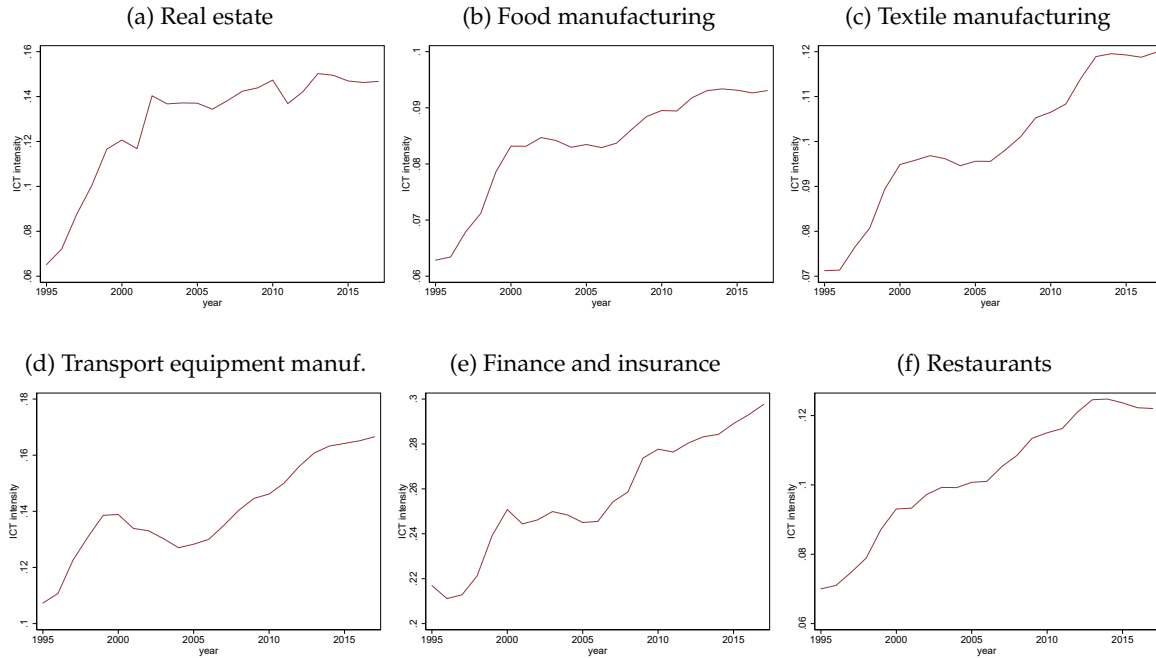
We group commodities into 25 broader categories and in Figure 2 plot the ICT intensity of six of them over time. In anticipation of the link to consumption categories in the CEX in the next section, we consider the six most important categories for consumption: Real estate, food manufacturing, textile manufacturing, transport, finance & insurance and restaurants.³ 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 not-

²All sectors except for the construction sector are matched. For 1996-2004 we use the 1997 matrix for 2005-2017 we use the 2012 matrix.

³The following I-O commodities are in these categories: real estate: 5310HS; food manufacturing: 311*, 312*; textile manufacturing: 313*, 314*; transport: 336*; finance and insurance: 52*; restaurants: 722*.

able finance and insurance has a much higher ICT intensity throughout the whole sample period while 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 I-O commodities. The weights correspond to their share in value added. *Source:* BEA and own calculations.

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

We observe only total expenditure, 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. But generally, due to the under-reporting by the rich, the CEX will rather understate consumption inequality. The theoretical part of this paper will offer the opportunity to cleanly separate prices

and quantities consumed.

As explained in more detail in Appendix A.1, we divide households into ten equal-sized bins based on labor income. For each income group, we create a weighted average of expenditure by year. Using the concordance table of Borusyak and Jaravel (2018), we link 809 CEX commodities to 159 I-O commodities. Each commodity in the CEX is matched to a unique I-O industry.⁴ Note that the concordance table treats housing as a service from the real estate sector rather than as a good from the construction sector.

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

| industry | 1996-99 | | | 2010-14 | | |
|-----------------------------------|---------|-------|-------------|---------|-------|-------------|
| | 1st | 10th | ratio | 1st | 10th | ratio |
| real estate (incl. construction) | 22.12 | 24.86 | 0.89 | 24.80 | 27.36 | 0.91 |
| food, beverages, tobacco | 17.19 | 8.79 | 1.96 | 16.03 | 7.33 | 2.19 |
| transport equipment | 8.12 | 9.15 | 0.89 | 8.26 | 10.45 | 0.79 |
| restaurants | 7.24 | 8.15 | 0.89 | 6.96 | 7.32 | 0.95 |
| textile, apparel, leather | 7.05 | 5.49 | 1.28 | 6.61 | 4.73 | 1.40 |
| chemicals, petroleum | 6.22 | 4.83 | 1.29 | 6.72 | 4.94 | 1.36 |
| finance,insurance | 5.20 | 7.32 | 0.71 | 5.15 | 6.85 | 0.75 |
| utilities | 4.88 | 2.97 | 1.64 | 5.19 | 3.02 | 1.72 |
| information | 4.33 | 3.59 | 1.21 | 4.49 | 4.04 | 1.11 |
| other services | 2.99 | 4.29 | 0.70 | 2.87 | 4.76 | 0.60 |
| misc. manufacturing | 2.94 | 3.80 | 0.77 | 2.87 | 4.76 | 0.60 |
| machinery,electronics,electronics | 2.45 | 2.92 | 0.84 | 1.95 | 2.60 | 0.75 |
| agriculture | 2.25 | 1.21 | 1.86 | 2.13 | 1.08 | 1.97 |
| health | 1.73 | 1.73 | 1.00 | 1.38 | 1.92 | 0.72 |
| trade,transport,warehousing | 0.89 | 1.68 | 0.53 | 0.96 | 1.51 | 0.64 |
| wood,furniture | 0.88 | 1.60 | 0.55 | 0.84 | 1.29 | 0.65 |
| professional services | 0.65 | 0.65 | 1.00 | 0.35 | 0.67 | 0.52 |
| rental and leasing | 0.64 | 1.54 | 0.42 | 0.67 | 1.31 | 0.51 |
| paper, printing | 0.62 | 0.47 | 1.32 | 0.52 | 0.38 | 1.37 |
| education | 0.57 | 1.39 | 0.41 | 0.55 | 2.32 | 0.24 |
| rubber, nonmetallic minerals | 0.42 | 1.01 | 0.42 | 0.42 | 0.57 | 0.74 |
| admin support | 0.24 | 0.48 | 0.50 | 0.31 | 0.52 | 0.60 |
| arts, entertainment, recreation | 0.22 | 1.02 | 0.22 | 0.28 | 1.10 | 0.25 |
| accommodation | 0.15 | 0.85 | 0.18 | 0.14 | 0.86 | 0.16 |
| primary and fabricated metal | 0.02 | 0.04 | 0.50 | 0.01 | 0.03 | 0.33 |

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 labor income distribution as well as their ratio, both for 1996-99 and for 2010-2014. These are ranked by their importance for low-income households. The largest share of income – around 30% – is being spent on housing and therefore gets allocated to real estate. There is little difference across time or across income deciles. 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.

⁴In contrast, a commodity would often be matched to several consumption categories. Therefore we convert consumption categories into commodities rather than the other way round. These 159 I-O commodities are the ones that are relevant for consumption in the CEX.

Some categories are consumed more extensively by the rich, in particular accommodation, finance and insurance, entertainment and recreation, and education. 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 patterns over time.

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.

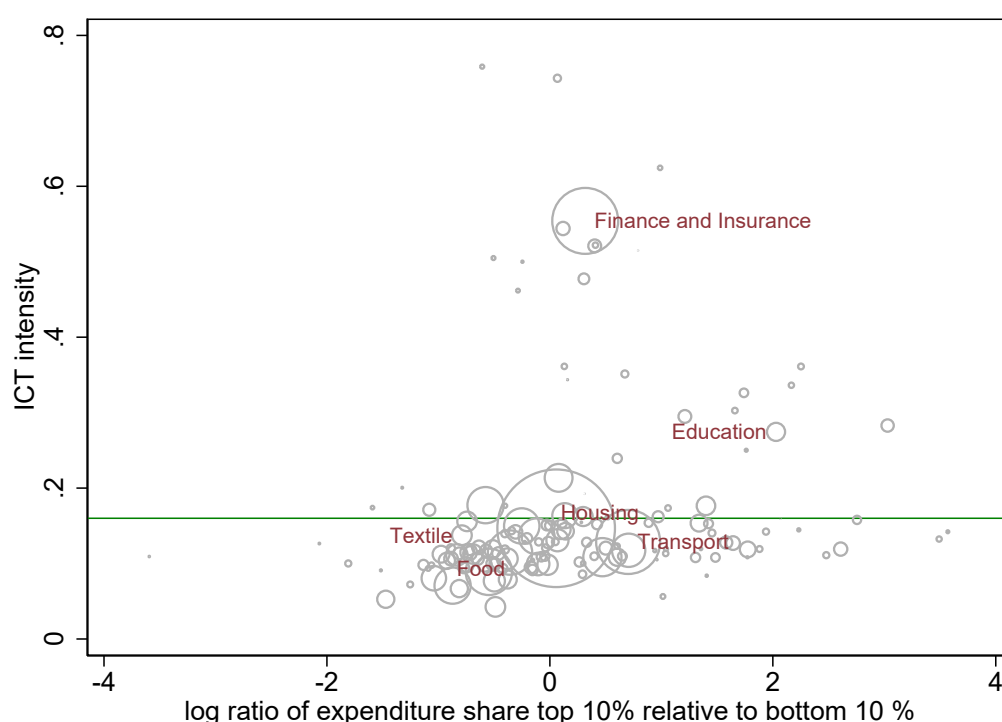
2.4 Digitalization share of consumption

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 over 159 final commodities in the year 2017⁵. Each dot corresponds to an I-O 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. 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.

⁵Data for 1997 and 2007 are provided in the Appendix, Figure A4.

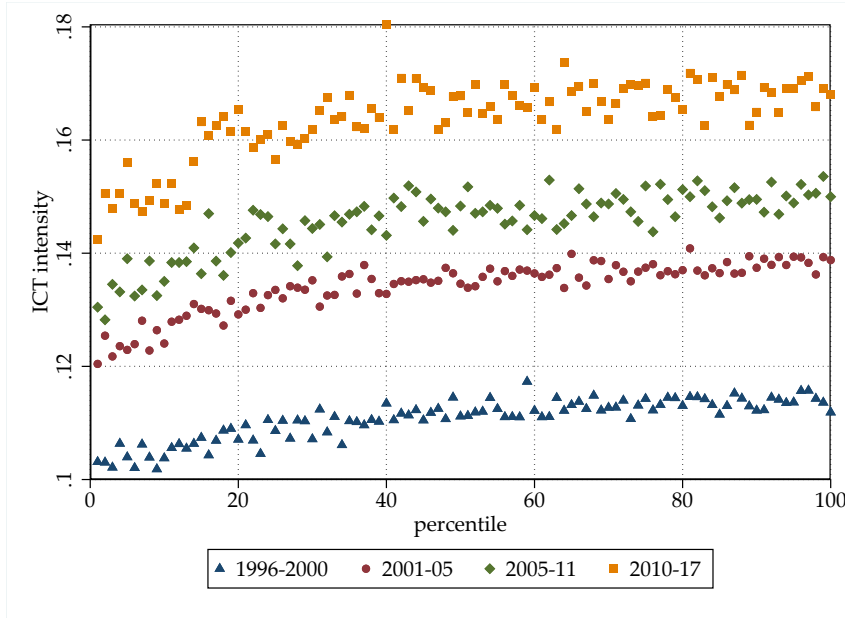
Figure 3: ICT intensity of vs. relative expenditure shares by commodity in 2017



Note: Each circle corresponds to one of the 159 I-O 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 is 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 decile. These reflect 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 affects the consumption bundle of the rich slightly more. The inequality in digital consumption has thus been increasing over the last 20 years. The pattern we find is similar to the one reported by Hubmer (2018), who shows that the average labor share intensity of the consumption basket increases with income. This aligns with our findings, as the labor-intensive goods also tend to be the more ICT-intensive goods.

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 year, 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. To quantify the effect, we need to carry out a more structured analysis. First, we need an understanding how commodity prices have changed depending on ICT intensity. The analysis is then followed by an introduction of a model framework in the following sections.

2.5 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, similar to Jaravel (2019). We match by hand 632 CEX product categories with 207 BLS price data series between 1996 and 2017.⁶ We then create aggregate price indices at the level of 159 different IO commodities using the Törnqvist price index and our CEX-IO concordance. 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 J individual goods, p_j , where the weights are nominal expenditure shares of the median household s_j ,

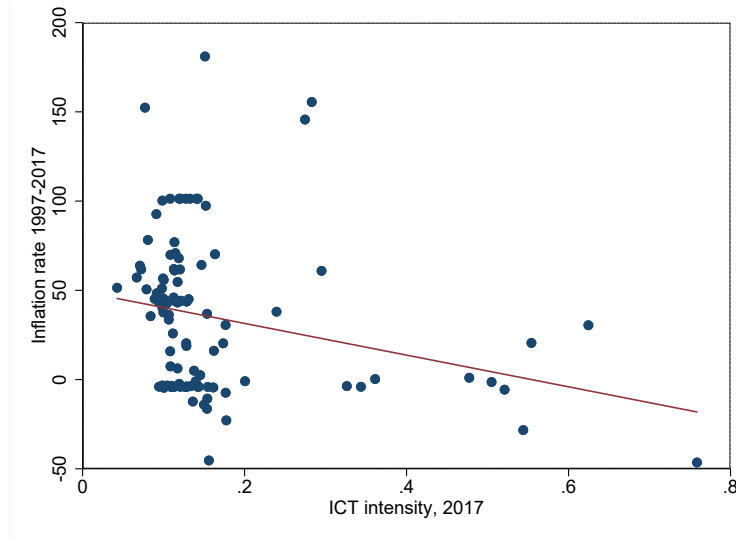
$$\Delta p_t = \sum_{j=1}^J \ln \left(\frac{p_{j,t}}{p_{j,t-1}} \right) \left(\frac{s_{j,t} + s_{j,t-1}}{2} \right)$$

⁶Some 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.

The level is recovered recursively relative to a base year 0 with $p_0 = 1$ as

$$p_t = p_{t-1} \exp(\Delta p_t)$$

Figure 5: Consumer price inflation and ICT intensity



Note: The graph plots percentage point changes in ICT intensity (post-IO) against inflation in consumer prices between 1997 and 2017. Each dot represents an I-O commodity. *Source:* BEA, CEX, BLS and own calculations.

As Figure 5 shows, between 1997 and 2017 consumer price inflation has been lower for commodities that have a higher ICT intensity. Among the commodities with the largest price decline are software publishers (511200) and electronic computer manufacturing (334111), which have a high ICT intensity. (A large price decline was also observed in doll, toy and game manufacturing (339930), which has a low ICT intensity but possible a large share of imports.) The largest price increases are in low-ICT intensity commodities like office supply manufacturing (339940), water, sewage and other systems (221300) and 611100 and 611A00, which both refer to education. The results are similar when using shorter time periods. The coefficient of correlation is -0.21. Thus, consumer with a consumption basket tilted towards ICT-intensive goods fare better. 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.

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 technology, who each produce a 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 Gaggli (2018) in setting up the following nested production structure:

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

$$SW_i = \left[\gamma_i H_i^{\frac{\epsilon_i-1}{\epsilon_i}} + (1 - \gamma_i) IT_i^{\frac{\epsilon_i-1}{\epsilon_i}} \right]^{\frac{\epsilon_i}{\epsilon_i-1}} \quad (1)$$

with ϵ_i being the elasticity of substitution between high-skill labor and ICT capital. Values smaller than 1 indicate that both inputs tend to be complements. Values larger than 1 indicate that they are more substitutable.

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

$$TW_i = \left[\phi_i L_i^{\frac{\eta_i-1}{\eta_i}} + (1 - \phi_i) SW_i^{\frac{\eta_i-1}{\eta_i}} \right]^{\frac{\eta_i}{\eta_i-1}} \quad (2)$$

with η_i as the elasticity of substitution between low-skill labor and the skilled work composite. If η_i is larger than 1, then low skill labor tends to be more substitutable to skilled work.

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 Gaggli, 2018 and Koh et al., 2020):

$$Y_i = K_i^{\alpha_i} TW_i^{1-\alpha_i} \quad (3)$$

There is perfect competition in each sector. First-order conditions yield the following equations for factor returns

$$K_i : \quad p_i \alpha_i \frac{Y_i}{K_i} = r_K \quad (4)$$

$$L_i : \quad p_i (1 - \alpha_i) \phi_i \frac{Y_i}{TW_i^{\frac{\eta_i-1}{\eta_i}} L_i^{\frac{1}{\eta_i}}} = w_L \quad (5)$$

$$H_i : \quad p_i (1 - \alpha_i) (1 - \phi_i) \gamma_i \frac{Y_i SW_i^{\frac{\eta_i-1}{\eta_i}}}{SW_i^{\frac{\epsilon_i-1}{\epsilon_i}} TW_i^{\frac{\eta_i-1}{\eta_i}} H_i^{\frac{1}{\epsilon_i}}} = w_H \quad (6)$$

$$IT_i : \quad p_i(1 - \alpha_i)(1 - \phi_i)(1 - \gamma_i) \frac{Y_i SW_i^{\frac{\eta_i - 1}{\eta_i}}}{SW_i^{\frac{\epsilon_i - 1}{\epsilon_i}} TW_i^{\frac{\eta_i - 1}{\eta_i}} IT_i^{\frac{1}{\epsilon_i}}} = r_{IT} \quad (7)$$

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 . The output of sector 2 can be used for consumption as well as to produce 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

$$Y_1 = C_1 + I_K \quad Y_2 = C_2 + \mu I_{IT} \quad (8)$$

μ is the relative price of ICT capital and at the same time measures progress in ICT technology (see e.g. Karabarbounis and Neiman, 2014; Eden and Gaggli, 2018 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.⁷

Capital follows the standard law of motion

$$K' = (1 - \delta_K)K + I_K \quad IT' = (1 - \delta_{IT})IT + I_{IT} \quad (9)$$

where δ_K and δ_{IT} 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. As in Herrendorf et al. (2013), the period utility function takes the form

$$u_j = \ln \left(\left[(1 - \omega) (C_{1,j} - \bar{C}_1)^{\frac{\sigma-1}{\sigma}} + \omega C_{2,j}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \right) \equiv \ln(C_j) \quad (10)$$

where \bar{C}_1 can be either larger than zero, making good 1 the necessity good, or smaller than zero, making it the luxury good. While we do not impose any restrictions at this point in the model, our empirical analysis has shown that the less ICT intensive good 1 is the necessity good. This implies that poor households (the low-skill households) have a larger share of good 1 in their basket than rich high-skill households. If σ is relatively small, then good 1 and 2 are partial complements to

⁷In 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.

the consumer. If σ is large then they tend to be substitutes⁸. We are going to use our data regarding expenditure shares for consumption goods to discipline the parameters of the utility function.

The high-skill households jointly provide \bar{H} units of labor, the low-skill households \bar{L} . 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.4.⁹ 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

$$C_{1,H}P_1 + C_{2,H}P_2 + I_KP_1 + I_{IT}P_2 = \bar{H}w_H + Kr_K + ITr_{IT} \quad (11)$$

L households are excluded from capital markets and are hand-to-mouth consumers. The budget constraint for type L households is

$$C_{1,L}P_1 + C_{2,L}P_2 = \bar{L}w_L \quad (12)$$

Both types solve an intratemporal optimization problem, deciding between consumption of goods 1 and 2. Define the aggregate price index

$$P = \left[(1 - \omega)^\sigma P_1^{1-\sigma} + \omega^\sigma P_2^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \equiv 1 \quad (13)$$

which we normalize to 1, such that wages, returns and the prices of the two goods are measured in units of the overall price level in the economy.

Solving via expenditure minimization, the Lagrangian of household j is

$$\mathcal{L} = P_1C_{1,j} + P_2C_{2,j} - \lambda \left(\left[(1 - \omega) (C_{1,j} - \bar{C}_1)^{\frac{\sigma-1}{\sigma}} + \omega C_{2,j}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} - U_j \right) \quad (14)$$

and results in the following Hicksian demand functions

$$C_{1,j} = C_j \left(\frac{P}{P_1} \right)^\sigma (1 - \omega)^\sigma + \bar{C}_1 \quad C_{2,j} = C_j \left(\frac{P}{P_2} \right)^\sigma \omega^\sigma \quad (15)$$

The L-type households cannot save, so they consume all of their income in each period. The H-type households also solve an intertemporal optimization problem, choosing consumption vs. investment in the two sectors. We can express the Bellman equation as

$$V(K, IT) = \max(\ln(C_H)) + \beta \mathbb{E} [V(K', IT')] \quad (16)$$

which we solve subject to the household budget constraint eq (11). This results in the Euler equa-

⁸For $\bar{C} = 0$ and $\sigma = 1$ the consumption aggregator collapses to the familiar Cobb Douglas function with share ω for good 2.

⁹Jaimovich et al. (2020) make the same assumption about capital ownership.

tions

$$\frac{P_2}{C_{HP}} = \beta \frac{P'_2}{C'_{HP'}} \left(\frac{\frac{R'_{IT}}{P'_2} + \mu'(1 - \delta'_{IT})}{\mu} \right) \quad (17)$$

$$\frac{P_1}{C_{HP}} = \beta \frac{P'_1}{C'_{HP'}} \left(\frac{R'_K}{P'_1} + (1 - \delta_K) \right) \quad (18)$$

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

$$\frac{P'_2}{P_2} \left(\frac{\frac{R'_{IT}}{P'_2} + \mu'(1 - \delta'_{IT})}{\mu} \right) = \frac{P'_1}{P_1} \left(\frac{R'_K}{P'_1} + (1 - \delta_K) \right) \quad (19)$$

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

3.4 Market clearing

The two capital markets clear

$$K = K_1 + K_2 \quad IT = IT_1 + IT_2 \quad (20)$$

The markets for consumption goods clear

$$C_1 = C_{1,L} + C_{1,H} \quad C_2 = C_{2,L} + C_{2,H} \quad (21)$$

The labor markets clear

$$\bar{L} = L_1 + L_2 \quad \bar{H} = H_1 + H_2 \quad (22)$$

The resource constraint eq. (8) is satisfied as well.

4 Calibration and Simulation

This section calibrates the model. We are not after assessing how much digitalization increases the skill premium or which effect it has on relative prices. Instead, we assume that digitalization can fully explain the changes observed in these two variables, and we aim to find out how digitalization affects consumption of each skill group, and how strong the price and income channels are.

We construct a series of μ , the relative price of ICT capital, from the data. This represents our exogenous digitalization shock, which we are going to feed into the model. The change in μ is the only time-varying element. Then we calibrate the parameters in the production function and the

preference parameters to match key moments from the data.

In the simulation, we assume that the economy is at a pre-digitalization steady state in 1960. Between 1960 and 2017, digitalization takes place. 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.

4.1 Industry classification

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 industries¹⁰ are classified to be part of the ICT-intensive sector that corresponds to our sector 2. Here and in other parts of the calibration, we choose the years around 1996 because this is the first year 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%.

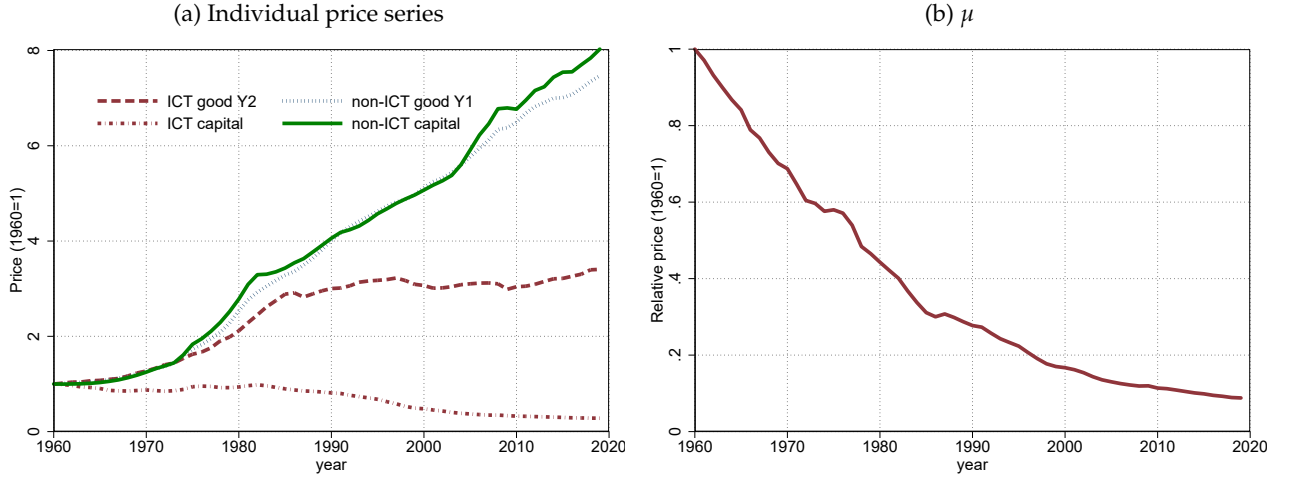
Our model assumptions imply that those industries clustered in the ICT-intensive sector are the only producers of ICT capital, while at the same time they do not produce non-ICT capital. As industries in this sector are producers of ICT goods (such as computer and electronic products, computer systems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

4.2 Advances in Digital Technology

According to eq. (8), 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 6 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. 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_{IT} relative to Y_2 . The resulting μ is shown in panel (b). We will feed a smoothed version of this series in Figure 6 (b) into our simulation as an exogenous change in ICT technology.

¹⁰These 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 6: Prices of final consumption goods and capital



Note: Panel (a) shows Törnqvist price indices based on industry- and asset-level data. Industries are classified into ICT- and non-ICT sector according to the clustering described in section 4.1. Assets are classified according to the definition used in Section 2.1. μ in panel (b) is the price series for ICT capital (red dashed-dotted line of panel (a)) relative to the price series for good 2 (red dashed line). *Source:* NIPA and own calculations.

4.3 Other parameters

The model has a large number of parameters that need to be calibrated. There are three parameters in the utility function, σ , ω and \bar{C} . In addition, ten parameters in the production function: Four elasticities of substitution ($\epsilon_i, \eta_i \forall i = 1, 2$) and six weights ($\alpha_i, \gamma_i, \phi_i$). Furthermore there are 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.¹¹ 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 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 nine 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. We target the following nine data moments: the

¹¹See Klump and de La Grandville (2000) for a more general discussion on how weights and elasticities are related.

percentage changes in the skill premium, the relative price, the ICT intensity, the labor share, the rate of return to ICT and non-ICT capital and the value-added share in sector 2 between 1960 and 2017, and the expenditure share of ICT goods in the consumption basket of high- and low-skill households in 2017.¹² Given our focus on the price vs. income channel, we require the algorithm to match the declines in the relative price and the skill premium exactly. With nine parameters and nine data moments, the system is exactly identified under the assumption that the moments are orthogonal.

Table 2 shows the calibration fit. We match most data moments very well, in particular those that matter for consumption inequality, the expenditure share of the ICT-intensive good 2 in the budget of high-skill and low-skill households. The model can also generate the strong increase in the importance of ICT capital for both sectors and the stronger growth in value-added of sector 2 relative to sector 1. The model does less well in matching the decline in the labor share, which is likely due to the fact that we do not allow for changes in the supply of labor. However, since this is not the focus of our paper, we prefer μ to be the only time-varying element that drives the model. Also, the increase in the rate of return to non-ICT capital is predicted to be much larger in the model than in the data.

Table 2: Calibration fit

| Data moment | Source | Data | Model |
|---|------------|--------|--------|
| Pct. change in skill premium, 1960-2017 | ACS | 22.3 | 22.3 |
| Pct. change in relative price, 1960-2017 | NIPA | 134.6 | 134.6 |
| Expenditure share of good 2 in H budget, 2017 (%) | CEX | 17.8 | 18.0 |
| Expenditure share of good 2 in L budget, 2017 (%) | CEX | 15.6 | 15.5 |
| Pct. change in ICT intensity, 1960-2017 | BEA | 2096.2 | 2191.4 |
| Pct. change in value-added share of sector 2, 1960-2017 | BEA | 223.7 | 195.4 |
| Pct. change in labor share, 1960-2017 | EHS (2013) | -11.7 | -1.3 |
| Pct. change, rate of return to ICT capital, 1960-2017 | KN (2019) | -86.0 | -91.4 |
| Pct. change, rate of return to K, 1960-2017 | KN (2019) | 44.5 | 136.7 |

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.

Table 3 summarizes the calibrated parameters. First, consider the preference parameters. Good 2 has a lower weight in the consumption basket, which matches the fact that the expenditure share of this good is relatively low (although increasing over time) in the CEX data for all income deciles. Our calibration suggests that there is a large degree of substitutability between ICT-intensive and non-intensive goods as σ is substantially above one. Remember that the ICT intensity of the final commodities in Section 2.2 was measured taking input-output linkages into account, which means that each good in the data incorporates both ICT-intensive and non-ICT intensive intermediates. As a result, we can have two products that are relatively similar, but where one is classified as

¹²Consumption expenditure data starts in 1996. For the quantification of the price and income channels, it makes little difference if we target the value in 2017 or one of the earlier years, as it just results in a level shift of the series for both high- and low-income households but does not have a significant effect on the difference between the two series.

Table 3: Calibrated parameters

| Symbol | Value | Description | Source |
|---------------------------|-------|-----------------------------------|------------------|
| Preferences | | | |
| ω | 0.35 | Weight for ICT good 2 | estimated by SMM |
| \bar{C}_1 | 42 | Subsistence consumption | estimated by SMM |
| σ | 2.2 | El. of subst. ICT non-ICT goods | estimated by SMM |
| Skilled Work | | | |
| γ_1 | 0.97 | Weight high-skill inner nest in 1 | estimated by SMM |
| γ_2 | 0.25 | Weight high-skill inner nest in 2 | estimated by SMM |
| $\epsilon_1 = \epsilon_2$ | 0.78 | El. of Subst between H and ICT | estimated by SMM |
| Total Work | | | |
| ϕ_1 | 0.5 | Weight low-skill middle nest in 1 | estimated by SMM |
| ϕ_2 | 0.59 | Weight low-skill middle nest in 2 | estimated by SMM |
| $\eta_1 = \eta_2$ | 3.6 | El. of Subst betw L and SW in 1 | estimated by SMM |
| Final Production | | | |
| α_1 | 0.55 | Capital share in 1 | BEA |
| α_2 | 0.28 | Capital share in 2 | BEA |
| δ_{ICT} | 0.15 | Depreciation rate ICT capital | BEA |
| δ_K | 0.08 | 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 | 0.965 | Discount factor | |

ICT-intensive and the other not. Due to the substitutability, households will react to the strong price decline in the ICT-intensive good by substantially increasing the expenditure share of this good over time. \bar{C} steers the difference in consumption shares of the ICT-intensive good between high-skill and low-skill households. A larger \bar{C} increases the difference in consumption patterns between the two skill groups. It also determines how strongly households' consumption baskets change in response to increases in income (leaving aside price changes).

Regarding the parameters of the production function, we see not surprisingly that $\gamma_1 > \gamma_2$ and $\phi_2 > \phi_1$, as sector 2 is more ICT-intensive. The large elasticity of substitution between low-skilled labor and SW, η , suggests that low-skilled labor and ICT capital are strong substitutes, whereas ICT capital and high-skilled labor are complements through a low 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 Gaggli, 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 incomes, prices and consumption of high- and low-skill households. To quantify the drivers of consumption inequality, we carry out two counterfactual analyses, a compensatory variation and a decomposition exercise.

5.1 Income

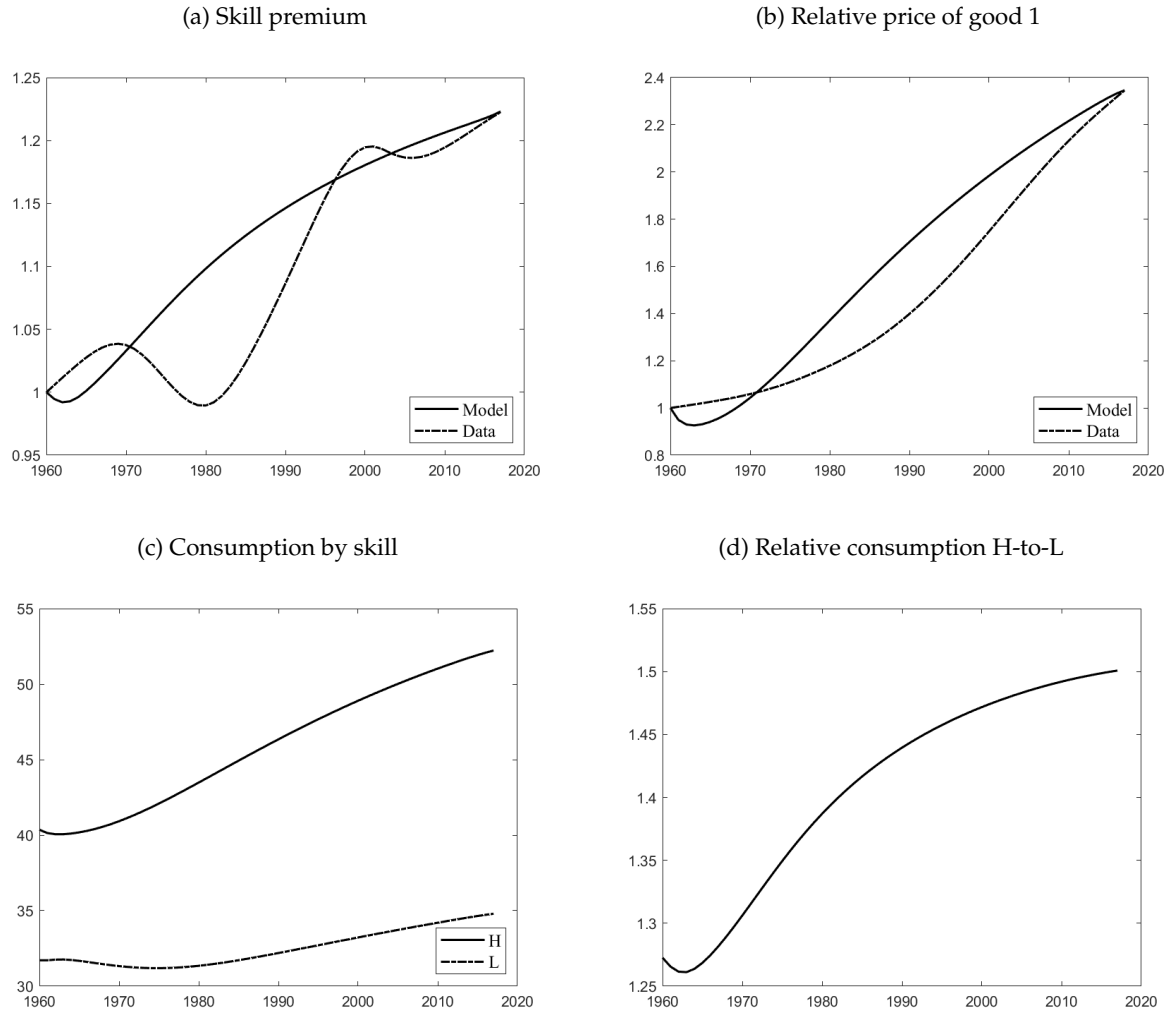
Panel (a) of Figure 7 shows the evolution of the skill premium W_H/W_L , normalized to 1 in 1960. The skill premium increases substantially both in the data and in the model. We have chosen the parameters to exactly match the long-run increase of the skill premium. The model does not capture some of 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. Due to the complementarity between ICT capital and high-skill labor, demand for high-skill labor rises. Simply put, if a firm uses more ICT capital such a 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 be reflected entirely in the change of wages. Let us consider separately the evolution of W_H and W_L . Due to the normalization $P = 1$, these are real values. High-skill workers benefit strongly from digitalization. Their real wage W_H rises by 37.6%. 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 12.5%, 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 24.7% over the simulation period. This number is considerably lower than the increase in the wage premium for two reasons: First, capital income grows by less than high-skill wage income. Second, the investment share increases. The share

Figure 7: 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. Panel (c) shows consumption of high-skill and low-skill households. Panel (d) shows consumption of high-skill households relative to low-skill households. *Data sources:* ACS and NIPA.

of wage income in net income increases slightly over time. Given these numbers, the increase in income inequality between high- and low-skill households is primarily driven by the increase in the skill premium and not by larger capital income.

5.2 Relative prices and relative demand

Panel (b) of Figure 7 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 for 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. While ICT investment was only 0.9% of total investment in 1960, it has risen to 10.6% by 2017.

Consumption demand also shifts towards good 2. Expenditure for good 2 in total consumption expenditures increases from 7.3% in 1960 to 16.9% in 2017. This is not solely due to the change in the relative price, but it is also linked to non-homothetic preferences¹³. 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 process is more important and therefore the relative price of good 1 still increases substantially.

\bar{C}_1 is calibrated to 42. In 1960, the ratio \bar{C}_1/C_1 is 0.37 for high-skill households and 0.43 for low-skill households, which means that subsistence consumption plays quite an important role. So it might seem straightforward to conclude that non-homothetic preferences are a determinant force in the evolution of consumption inequality. To assess how much they really matter, we carry out a simple experiment where we set $\bar{C}_1 = 0$, but keep the calibration the same otherwise. In this case, the expenditure share of good 2 is obviously higher both in 1960 and 2017. Furthermore, there are no differences in the expenditure shares between households. But importantly, the change in expenditure shares over time is very similar in the two settings: the increase of the expenditure share of good 2 is 131.5% in the case with non-homothetic preferences, compared to 127.0% with homothetic preferences. This means that relative price changes alone can generate a large change in consumption behavior. However, when focusing separately on high-skill households, the increase in the expenditure share of good 2 is 138.9% under non-homothetic preferences, so there the increase in disposable income plays a role. In contrast, for low-skill households, there is almost no difference in expenditure share growth between the two set-ups, because their real income does not grow much. We discuss a model version with $\bar{C}_1 = 1$ in more detail in Appendix B.

¹³The debate about whether it is non-homothetic preferences or relative prices that drive long-run changes in sectoral consumption shares plays a prominent role in the structural change literature. Depending on the model framework, this literature finds that either one or the other matters more in explaining the change from agriculture towards manufacturing towards services in the consumption baskets of households. See Herrendorf et al. (2013) for an overview and assessment. We are less interested in explaining consumption shares of different commodity types over time, and have set up the utility function such that a priori, it allows for both channels, with the calibration showing the parameter specification that can best explain the data.

5.3 Consumption inequality

Now we turn to changes in consumption inequality. Panel (c) of Figure 7 shows total consumption C by skill type. High-skill households have a higher level of consumption, and thus utility, throughout the whole simulation period, but their consumption also increases more over time. As shown in panel (d), initially, high-skill households have a 1.27 times larger consumption composite than low-skill households. This number increases to around 1.5 by the end of the simulation period, which is 18.1% higher. So digitalization has a substantial effect on consumption inequality.

Several recent papers have documented an increase in consumption inequality in the U.S. over the last decades¹⁴. Our results align with the findings of that literature. For example, Aguiar and Bils (2015) find that consumption inequality has increased by around 30% between 1980 and 2010 when accounting for measurement errors in the CEX. The measures are not directly comparable to ours as they consider a different time horizon and use a different consumption concept, but nevertheless point qualitatively to the same direction. 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.

5.4 Quantifying the importance of each channel

Both the income and the price channel work in favor of the high-skill households. But how important is each of them? This section aims to quantify the relevance of each channel via two different counterfactual exercises. Our first exercise is a compensatory variation, where we ask how much additional income we need to give to households in 2017 in order to compensate them for the increase in the relative price, and compare this to the actual income increase. We can thus compute in monetary units how much more high-skill households benefit from digitalization than low-skill households. The second exercise aims to quantify how much the price channel contributes to consumption inequality, by comparing consumption in the baseline model to a version with fixed relative prices.

For both approaches, we specify demand for goods 1 and 2 of household j as a function of disposable income Inc_j/P_2 and the relative price P_1/P_2 ,

$$C_{1j} = \left(\frac{Inc_j}{P_2} + \bar{C}_1 \left(\frac{\omega}{1-\omega} \right)^\sigma \left(\frac{P_1}{P_2} \right)^\sigma \right) / \left(\frac{P_1}{P_2} + \left(\frac{\omega}{1-\omega} \right)^\sigma \left(\frac{P_1^t}{P_2} \right)^\sigma \right)$$

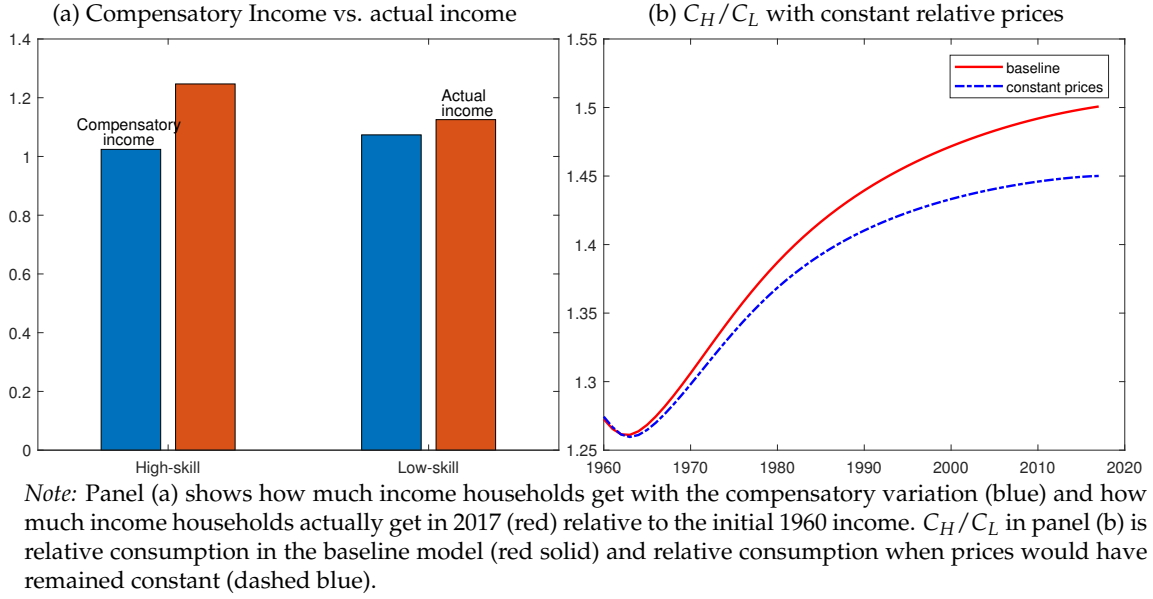
$$C_{2j} = \frac{Inc_j}{P_2} - \left(\frac{P_1}{P_2} \right) C_{1j}$$

The income of low-skill households is simply their wage income. High-skill households earn both a wage income and a capital income, from which we deduct the investments in order to compute disposable income for consumption.¹⁵

¹⁴E.g. Aguiar and Bils (2015) and Attanasio and Pistaferri (2014); see Attanasio and Pistaferri (2016) for an overview. The discussion often centers on whether consumption inequality has mirrored income inequality.

¹⁵This exercise keeps investments fixed. In a more evolved set-up, investment decisions would also be affected by

Figure 8: Welfare Decomposition



For the compensatory variation, we consider the initial welfare level of households as of 1960. This equals the utility u_j from eq. (10), using the actual amounts consumed in 1960 under 1960 prices. Then, we assume that prices in 1960 were as in 2017. The compensatory income is the income needed to achieve the same level of utility as in 1960, u_j , using the 2017 prices.¹⁶ In the next step, we compare this compensatory income with the actual real increase in income that households experience between 1960 to 2017. If the actual increase in real income is larger than the compensatory real income, households are better off in the post-digitalization world. We do this exercise separately for high-skill and low-skill households.

Figure 8 (a) 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 income. The blue bar shows how much income households need in 2017 to achieve the same level of welfare as in 1960. This is the amount of income required to compensate them for the relative price change. As low-skill households rely more on consumption of good 1 than high-skill households, their compensatory income is slightly larger (107.4% of income in 1960, compared to 102.4 % for high-skill households). The red bar shows how much more real income households actually receive in 2017 relative to 1960. As digitalization increases wages of high-skill households more, their disposable income increases by 24.7 %. In contrast, low-skill household's real income increases by 12.5%. 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 welfare gain 21.8% ($1.247/1.024-1$), whereas low-skill households gain only 4.8%. 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.

changes in prices. We abstract from this adjustment margin in order to facilitate the comparison between high-skill and low-skill households.

¹⁶Note that the aggregate price index is the same for both households and approximates an economy wide CPI. This means that the weights of the goods do not exactly correspond to the expenditure shares of high- or low-skill households. It is a general weakness of the CPI that it does not correspond to price indices faced by individuals

Second, for the counterfactual analysis of consumption, we keep relative prices fixed as in 1960. We then compute consumption in a setting where prices are fixed, but (real) incomes increase over time for each skill type. As before, C_H/C_L is our measure of consumption inequality. This analysis measures the strength of the income channel as a driver of inequality. By comparing this counterfactual version with the baseline in which relative prices change, we obtain the price effect.

Figure 8 (b) shows how consumption would have evolved under fixed relative prices. Without the relative price change that favored rich households more, consumption inequality would have increased by 22.5% less. Including changes in relative prices, consumption inequality increases by 18.1%, while in the absence of the price effect the increase is only 14.0%. 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.

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 final commodity producers. Combining this measure with consumption data along the income distribution, we show that rich households have a larger digitalization share than poor households. Furthermore we provide evidence that the price change of digital relative to non-digital 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 consumption inequality has increased by 18%. 22.5% of this increase is attributed to changes in relative prices. These results point out that the effect of digitalization on consumption inequality could be 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 D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, Volume 4, pp. 1043–1171. Elsevier.
- Acemoglu, D. and P. Restrepo (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review* 108(6), 1488–1542.
- Adão, R., M. Beraja, and N. Pandalai-Nayar (2020). Technological transitions with skill heterogeneity across generations. *NBER Working Paper No. 26625*.
- Aghion, P., C. Antonin, X. Jaravel, and S. Bunel (2020). What are the labor and product market effects of automation? New evidence from France. *CEPR Discussion Paper 14443*.

- Aguiar, M. and M. Bils (2015). Has consumption inequality mirrored income inequality? *American Economic Review* 105(9), 2725–2756.
- Attanasio, O. P. and L. Pistaferri (2014). Consumption inequality over the last half century: some evidence using the new PSID consumption measure. *American Economic Review* 104(5), 122–26.
- Attanasio, O. P. and L. Pistaferri (2016). Consumption inequality. *Journal of Economic Perspectives* 30(2), 3–28.
- Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review* 103(5), 1553–1597.
- Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in US wage inequality: Revising the revisionists. *Review of Economics and Statistics* 90(2), 300–323.
- 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.
- Borusyak, K. and X. Jaravel (2018). The Distributional Effects of Trade: Theory and Evidence from the United States. Technical report, Available at SSRN, no. 3269579.
- 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.
- 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.
- Graetz, G. and G. Michaels (2018). Robots at work. *Review of Economics and Statistics* 100(5), 753–768.
- Hémous, D. and M. Olsen (2020). The rise of the machines: Automation, horizontal innovation and income inequality. *American Economic Journal: Macroeconomics* (forthcoming).
- Herrendorf, B., R. Rogerson, and Á. Valentinyi (2013). Two perspectives on preferences and structural transformation. *American Economic Review* 103(7), 2752–2789.
- Herrendorf, B., R. Rogerson, and Á. Valentinyi (2014). Growth and structural transformation. In *Handbook of Economic Growth*, Volume 2, pp. 855–941. Elsevier.
- Horowitz, 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.

- 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. Santaaulalia-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. (2018). Automation, growth and factor shares. *Unpublished manuscript*.
- McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 995–1026.
- Moll, B., L. Rachel, and P. Restrepo (2021). Uneven growth: Automation’s impact on income and wealth inequality. *NBER Working Paper No.21595*.
- 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

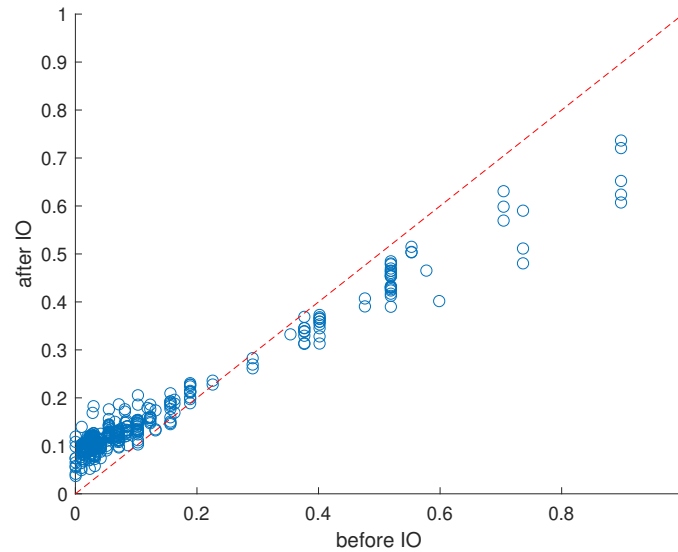
We use the BEA’s Input Output Accounts and focus on the detailed Input-Output tables after re-definitions. We use producer-value tables, which means that the distribution margin (i.e. the cost of wholesaling, retailing and transportation) is modeled as a flow from the distribution industries to the final 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 Horowitz 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 I-O structure



Note: The graph shows ICT intensities for 2012 before the I-O structure has been taken into account (x-axis) and afterwards (y-axis). Each dot represents an I-O 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. Our income measure is gross labor income, which is captured by the variable FSALARYX (FSALARYM in later vintages) in the interview survey and FWAGEX (FWAGEXM) in the diary survey.¹⁷ 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.

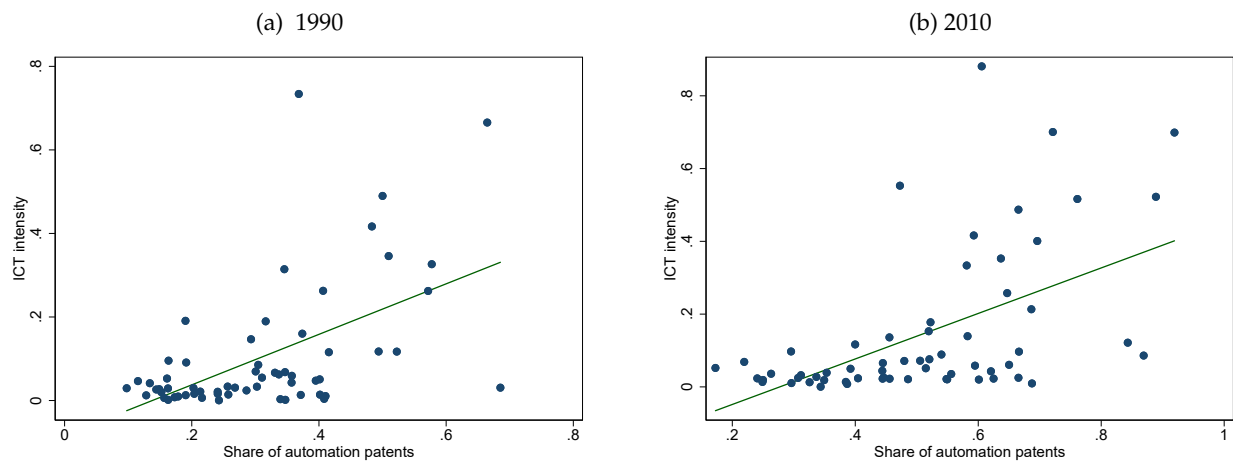
¹⁷An alternative would be to consider total before-tax household income, which corresponds to FINCBTAX (FINCBTXM in later vintages) for the interview survey and FINCBFX (FINCBFXM) for the diary survey. In addition to labor income, these broader measures include farm and non-farm business income, social security income, interest on savings accounts or bonds, income from dividends, royalties, estates and trusts and rental income. However, 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).

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. An additional problem is that the CEX only captures about 70% of total household expenditure. But other surveys like the PSID are not detailed enough and do not offer a long-enough time series dimension.

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 2010¹⁸. 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.

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

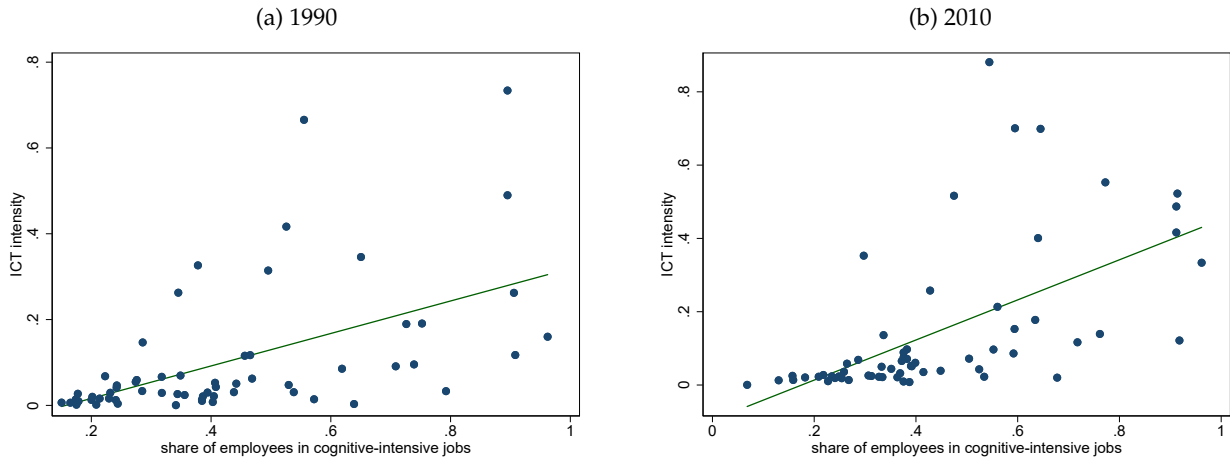


Note: The graph plots ICT intensity against the share of automation patents in the total number of patents by BEA industry. Each dot represents one industry, the line shows a linear prediction with 95% confidence interval. The left panel shows raw data for 1990, the right panel for 2010. *Source:* Mann 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. Gaggli and Wright (2017) and

¹⁸The plots would look similar for other years.

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



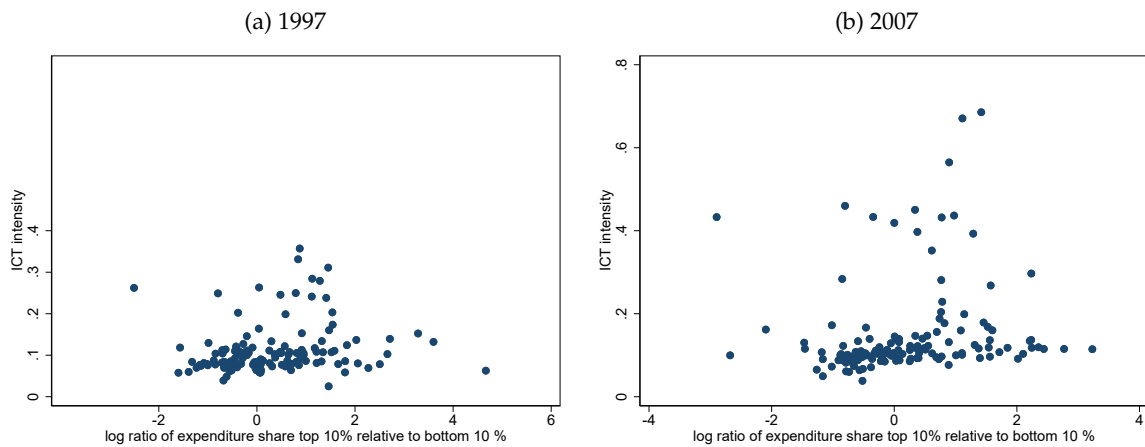
Note: The graph plots ICT intensity against the share of employees that work in cognitive-intensive jobs by BEA industry for 1990 (left panel) and 2010 (right panel). Each dot represents one industry, the line shows a linear prediction with 95% confidence interval. *Source:* Census, American Community Survey, Autor and Dorn (2013), BEA and own calculations.

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 A3 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.

A.3 Additional figures

Figure A4: ICT intensity of consumption, top 10% relative to bottom 90 %



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 with 95% confidence interval. *Source:* Census, American Community Survey, Autor and Dorn (2013), BEA and own calculations.

A.4 Income and asset data from the SCF

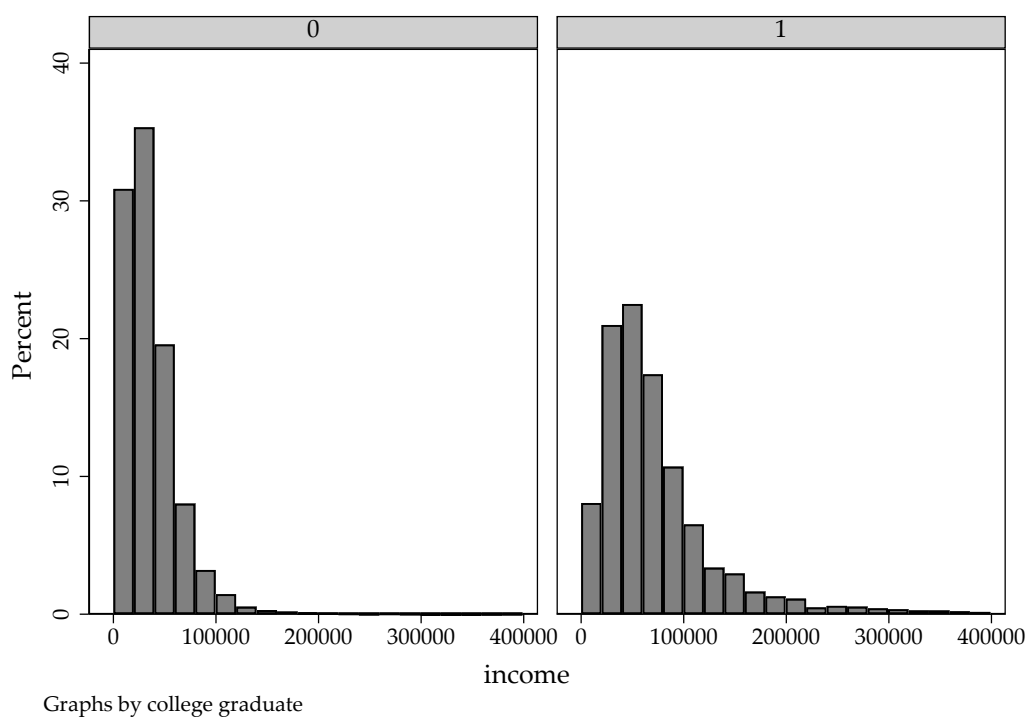
The aim of this section is to motivate our assumption that capital is owned only by high-skill households. To show that this assumption is supported by data, we use data from the Survey of Consumer Finances (SCF) between 1998 and 2013. We focus on households where the head is between 20 and 64 years 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 A1: 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 |

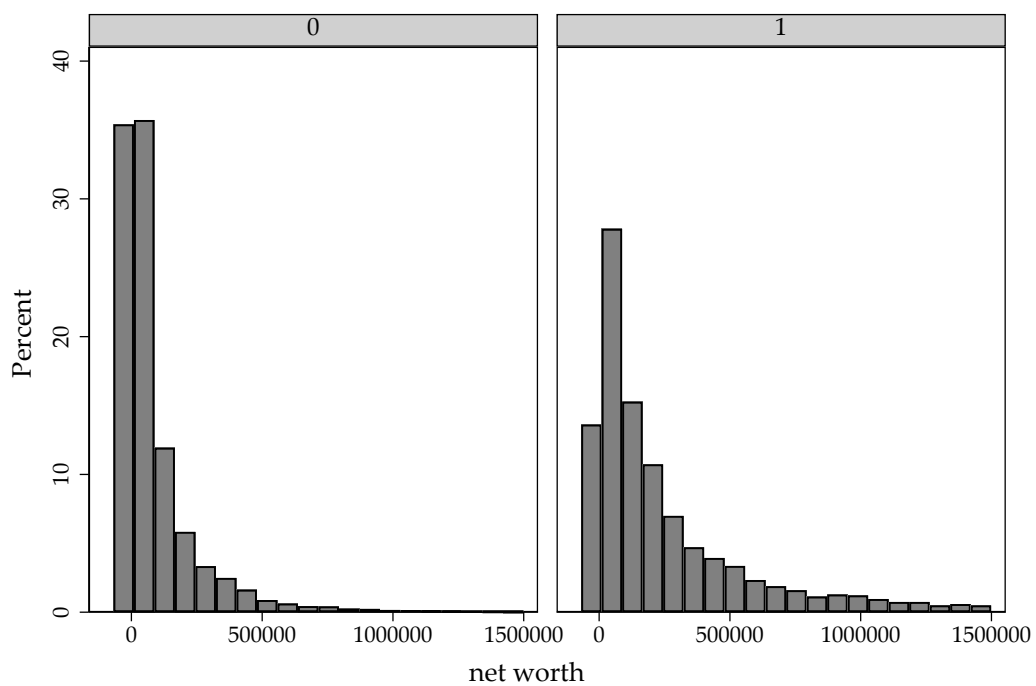
All numbers are in 2019 US Dollars and refer to the median household in each education group. Total income is the sum of labor income, capital income and transfer income.

Figure A5: 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 A6: Net wealth distribution by skill

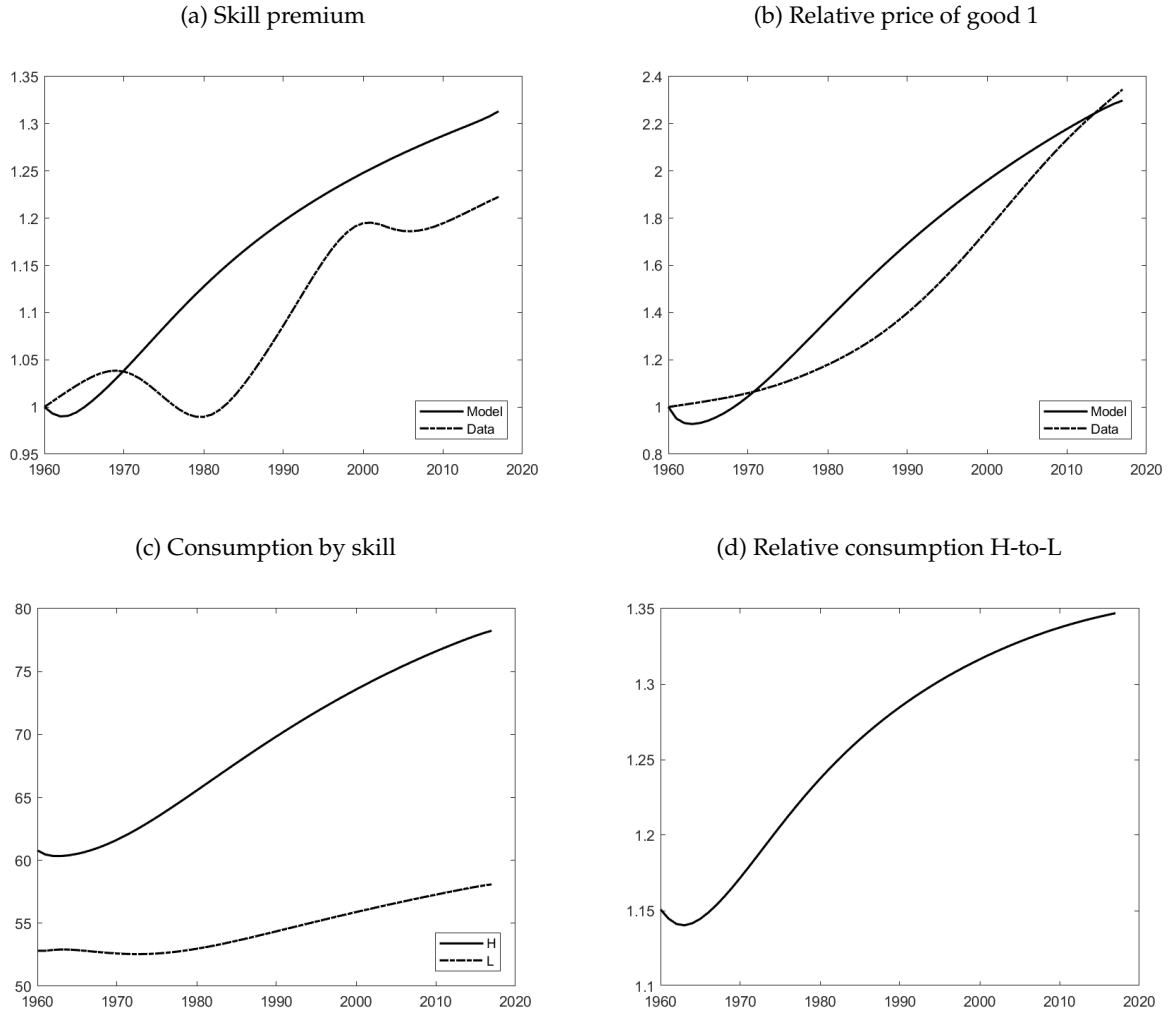


Graphs by college graduate

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

B Model Appendix

Figure B1: Simulation results with $\bar{C}_1 = 0$



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. Panel (c) shows consumption of high-skill and low-skill households. Panel (d) shows consumption of high-skill households relative to low-skill households. *Data sources:* ACS and NIPA.

We conduct a simulation with the same parameterization as in the model presented in the main text, except that we set $\bar{C}_1 = 0$ here. This means that we have homothetic preferences and both households spend the same share of their income on good 2. Figure B1 summarizes the simulation results analogous to the main text. Importantly, the skill premium rises more in a model with $\bar{C}_1 = 0$. The reason for the stronger increase is due to a larger increase in investment of ICT capital that makes high-skill work more valuable. Initially, households in a model with $\bar{C} = 0$ have an overall larger share of consumption of good 2. As final output is distributed between investment and consumption, the initial investment share of ICT capital is much lower than in a world with a positive \bar{C}_1 . The effect of digitalization with homothetic preferences is then much stronger, the ICT intensity of the *whole* economy increases by more than 2250%, compared to just 2190% in the baseline model. This explains the higher skill premium. The difference to the baseline simulation is small, but the relative price rises a bit less here. Several forces are at work here. First, the demand increase for consumption of good 2 is less strong with $\bar{C}_1 = 0$ as its share on expenditures stays constant. This means that prices of good 1 should rise more. Another important force here

however is again the investment channel: The overall growth of sector 2 is less strong with $\bar{C}_1 = 0$ as there is no additional increase of consumption shares due to higher income. More of the additional growth of ICT goes therefore into the first sector. This creates relatively more technological upgrades there and makes the relative price increase less pronounced. Those two effects are very similar in size, the investment channel being slightly bigger though.

Consumption inequality is lower with homothetic preferences. Again, two forces with opposing signs are at work here. First, income of high-skill households rises more, as their wage income increases more as well. This increases consumption inequality. However, both households have the same exposure toward good 2, this means both households benefit equally from the relative price change. When \bar{C}_1 is greater than 0, rich households benefit more because they have a higher share of good 2 than low skill households. Here, the price effect is more important and consumption inequality increases substantially less with $\bar{C}_1 = 0$.

Note that the intuition in a model where investment and capital plays a driving role of technological process is very different from the usual models used in the structural change literature. There, consumption demand plays the dominant role and determines the size of sectors. In our model, consumption also (partly) crowds out investment with important effects on sectoral size, relative prices and factor prices.