From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2020

Essays on Automation and Labor Markets

Humlum, Anders

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp015h73q011x

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Essays on Automation and Labor Markets

Anders Humlum

A dissertation
presented to the faculty
of Princeton University
in candidacy for the degree
of Doctor of Philosophy

Recommended for acceptance
by the Department of Economics
Adviser: Stephen J. Redding

September 2020

© Copyright by Anders Humlum, 2020.

All rights reserved.

Abstract

How do recent automation technologies affect labor markets? And what can policymakers do to alleviate adverse impacts on workers? In this collection of essays, I provide empirical evidence on how automation shapes labor markets.

In Chapter 1, I use administrative data that link workers, firms, and robots in Denmark to study the distributional impact of industrial robots. I structurally estimate a dynamic model of the firm that rationalizes how firms select into and reorganize production around robot adoption. To understand the macroeconomic implications of robot adoption, I embed the firm model into a general equilibrium framework that endogenizes the choice of workers to switch occupations. Using the general equilibrium model, I estimate that industrial robots have increased average real wages by 0.8 percent but with substantial distributional consequences. Production workers employed in manufacturing have lost 6 percent from the arrival of robots.

In Chapter 2, I provide causal evidence on the individual and market level impacts of retraining subsidies. I study a reform that unexpectedly and temporarily eliminated all subsidies for 40 percent of the retraining courses available for unemployed workers in Denmark. I find that training activity reacts strongly to subsidies, and that retraining courses are effective in helping workers find jobs.

In Chapter 3, co-authored with Bjørn Bjørnsson Meyer, we first rank college majors according to whether their graduates work in Artificial Intelligence (AI) firms. We then use a regression discontinuity design to estimate the causal effects of admitting students to more AI relevant college majors. We find substantial labor market returns to college degrees that are relevant to firms that produce AI.

Acknowledgements

This dissertation would not have been possible without the incredible support of my adviser, Stephen Redding. Steve has a selfless and conscientious dedication to the economics profession that is truly unmatched. His multiple rounds of thorough comments on this dissertation constitute the sharpest and most constructive intellectual feedback I have ever received. Bo Honoré has been an invaluable mentor to me. From choosing a dissertation topic to getting ready for the job market, Bo provided decisive guidance at critical crossroads of my PhD journey. A tip for future students: whenever Bo opens his mouth, do pay close attention. He is the most thoughtful person I know.

I am grateful to have learned labor economics from Alexandre Mas, whose creativity as an empirical researcher I can only admire. I feel lucky to have assembled such a diverse dissertation committee. Steve, Bo, and Alex came together from three different fields to advise me on a subject that wasn't any of their main topics of expertise. Yet, their mutual respect laid the ideal grounds for me to write an interdisciplinary dissertation.

I have learned tremendously from the Trade group at Princeton. Esteban Rossi-Hansberg's awe-inspiring intellect and impressive level of ambitions challenged me to rise to higher levels. Eduardo Morales' multi-faceted approach to economics, pushing the frontier of how to connect theory, data, and statistics, is a style of research that I would like to emu-

late.

Finally, during my time in Princeton, I have learned more from my classmates than from anyone else. Adrien Bilal, Yann Koby, Mauricio Matsumoto, Mathilde Le Moigne, Charly Porcher, and Fabian Trottner – thank you for forming me as a researcher and as a person.

To Maria, whose instinctive wisdom
and unconditional love was the
North Star of this journey.

Contents

Abstract **iii**

Acknowledgements **v**

1 Robot Adoption and Labor Market Dynamics **1**

2 Individual and Market-Level Effects of Retraining Subsidies: Evidence From a Natural Experiment **102**

3 Artificial Intelligence and College Majors **158**

Chapter 1

Robot Adoption and Labor Market Dynamics

Versions of this chapter were previously presented at the Center for Applied Microeconometrics (CAM) Workshop, the Danish Academic Economist in North America (DAEiNA) Annual Meetings, the Danish International Economics Workshop (DIEW), the Asian Dynamics Initiative (ADI) Conference, and in seminars at various institutions.

1 Introduction

The arrival of industrial robots in modern manufacturing is one of the most salient technological changes in recent decades. Defined as "automatically controlled, reprogrammable, multipurpose manipulators programmable in three or more axes" (ISO 8373), industrial robots were developed for car assembly in the 1990s but have since diffused widely in manufacturing. Today, robot adopters account for half of manufacturing sales, and adoption rates are accelerating. The potential labor displacing effects of industrial robots have received much public attention, culminating when the European Parliament voted in 2017 on a proposal to tax the use of robotics (Delvaux, 2016).

This paper asks who gains and who loses when industrial robots are adopted. To answer this question, I use administrative data that link workers, firms, and robots in Denmark. My first contribution is to combine event studies with a structural model that rationalizes how firms select into and reorganize production around robot adoption. I find that firms expand output by 20 percent but shrink their wage bill on production workers, such as assemblers and welders, by 20 percent when they adopt industrial robots. Firms' total wage bill increases 8 percent as labor demand shifts toward tech workers, such as skilled technicians, engineers, and researchers. I structurally estimate a dynamic model of the firm that matches these reduced-form responses to robot adoption, the observed size premium in the selection of firms into robot adoption, as well as the S-shape in robot diffusion over time.

To understand the macroeconomic implications of robot adoption, I embed the firm

model into a general equilibrium framework that endogenizes the dynamic option for workers to reallocate across occupations. The estimated general equilibrium model captures several indirect effects of industrial robots that are not identified in micro-level diff-in-diff designs. These indirect effects include the extent to which the expansion of robot adopters crowds out non-adopter firms in product and labor markets, as well as the ability of workers to reallocate across occupations in response to equilibrium wage pressures from robot diffusion.

Using the general equilibrium model, I estimate that industrial robots have increased average real wages by 0.8 percent, but with substantial distributional consequences. At the opposite ends of the spectrum, I find that production workers employed in manufacturing have lost 6 percent in real wages, while tech workers have gained 2.3 percent. I find that welfare losses from robots are concentrated on old production workers. Younger workers, with less specific skills and a long career ahead of them, benefit from the option value of switching into tech and other occupations whose premiums rise as robots diffuse in the economy.

Occupational reallocation in response to industrial robots can account for 25 percent of the fall in the employment share of production workers and 8 percent of the rise in the employment share of tech workers in Denmark since 1990. The adoption of industrial robots have thus been a driver of employment polarization (Autor and Dorn, 2013; Goos et al., 2014). Without these labor supply responses, I find that the real wage loss of production workers from robots would have been five times larger.

My findings highlight the importance of allowing for labor supply responses when evaluating the distributional impact of industrial robots. I use a dynamic occupational

choice model that represents the state of the art for studying labor market dynamics in response to trade liberalizations (Dix-Carneiro, 2014; Traiberman, 2019), and I estimate the barriers to occupational switching using observed worker transitions together with a conditional choice probability (CCP) estimator that controls for the unobserved continuation values of workers.

As a final counterfactual exercise, I evaluate the dynamic incidence of a robot tax. The undistorted equilibrium of the model is efficient (except for markups in product markets), but I use the estimated model to quantify the distributional implications of a robot tax and to evaluate its impact on aggregate economic activity. I find that a temporary robot tax can be an effective way to slow the diffusion of industrial robots. However, compared to a permanent tax of similar magnitude, a temporary tax creates larger welfare losses per dollar of revenue collected and a larger fraction of its deadweight burden falls on workers. These larger adoption elasticities and relative efficiency losses reflect the forward-looking nature of adoption whereby firms foresee that the temporary tax will expire and postpone adoption until then. Based on the estimated responses, I conclude that a robot tax is an ineffective and costly way to redistribute income to production workers in manufacturing.

Evaluating the counterfactuals above requires solving the firm and worker problems jointly, and I develop a fixed-point algorithm for solving the dynamic general equilibrium of this class of models. A key property of the general equilibrium model is that the firm and worker problems are separable conditional on the path of wages. This separable structure is highly useful in estimation and in simulation. First, it allows me to estimate the firm (worker) model without specifying the problem of the worker (firm) by simply

conditioning on the observed path of wages. Second, it breaks the curse of dimensionality wherein firm variables become states for the worker, and worker variables become states for the firm. The separable structure enables me to incorporate the rich firm and worker heterogeneity estimated in the micro data, and still be able to compute the general equilibrium featuring joint firm and worker dynamics.

To measure robot adoption at the firm level, I leverage the fact that almost all industrial robots used in Denmark are not actually produced in the country. In particular, once an imported robot crosses the country border, it is recorded by the customs authorities under the 6-digit product code *847950 Industrial Robots*. The customs records, which contain information on the timing and value of firm robot imports, offer a unique opportunity to study what happens when firms adopt industrial robots. I supplement the customs records with a representative robot adoption survey conducted by Statistics Denmark, and I validate that these micro data sources on robot adoption align with industry-level measures used in the prior literature (Acemoglu and Restrepo, 2020). By merging the firm robot adoptions to the Danish matched employer-employee data, I obtain a dataset with unusually rich information on both firms and workers that is ideally suited to studying the distributional impacts of industrial robots.

This paper is related to and builds on several literatures. The most immediately related work is a recent series of papers that have collected reduced-form evidence on how industrial robots affect firm performance and labor market outcomes (Acemoglu and Restrepo, 2020; Bessen et al., 2019; Graetz and Michaels, 2018; Koch et al., 2019). I complement this work with two key structural contributions. First, I estimate a model of firm robot adoption that allows me to interpret the new reduced-form evidence in terms of

structural primitives. Second, I embed the model into a general equilibrium framework, enabling me to extend the identified micro-level effects to quantify the macroeconomic impacts of industrial robots. The two-sided nature of the general equilibrium model allows me to connect evidence on firm (e.g., Koch et al. (2019)) and worker outcomes (e.g., Dauth et al. (2018)) of robotization.

The methodology developed in this paper builds heavily on the literature of dynamic discrete choice models. The robot adoption model draws on the Rust (1987) optimal stopping model, and the labor supply module follows closely a series of structural labor papers, including Dix-Carneiro (2014) and Traiberman (2019). In the structural estimation, I build on the work by Doraszelski and Jaumandreu (2018) on estimating production functions with endogenous technical change, and I apply the methods of Arcidiacono and Miller (2011) on conditional choice probability (CCP) estimation of dynamic discrete choice models.

The remainder of the paper is structured as follows. Section 2 describes the Danish data and collects stylized facts on firm robot adoption. Sections 3 and 4 develop and estimate a partial equilibrium model of firm robot adoption. Section 5 estimates the labor supply module. Section 6 unites the firm and worker blocks, and then uses the general equilibrium model to estimate the distributional impact of industrial robots and to evaluate the incidence of a robot tax. Section 7 concludes.

2 Data

I use register data that link workers, firms, and robots in the Danish economy from 1995 to 2015. The dataset is the product of merging the Danish matched employer-employee data with two new micro data sources on firm robot adoption. This linked dataset contains unusually rich information on both firms and workers, making it ideally suited to studying the distributional impacts of industrial robots. The matched worker-firm-robot panel data offer a unique opportunity to study what happens, at the micro level, when industrial robots are adopted. The data contain detailed occupational codes of workers, allowing me to study how firms substitute between labor tasks in production. The universal coverage of the Danish data is essential for estimating the general equilibrium environment that robot adopters operate in.

The firm data come from the Firm Statistics (FirmStat) Register, which covers the universe of private-sector firms from 1995 to 2016. FirmStat associates each firm with a unique identifier, and provides annual data on many of the firm's activities, such as sales, number of full-time employees, and industry affiliation. The data on workers and establishments come from the Integrated Database for Labor Market Research (IDA), which covers the entire Danish population. IDA associates each person with her unique identifier, and provides annual data on many individual characteristics such as income, hours, hourly wage, detailed occupation, education, and other sociodemographics. To match the firm and worker data, I draw on the Firm-Integrated Database for Labor Market Research (FIDA), which links every firm in FirmStat with every worker in IDA who is employed

by that firm in week 48 of each year. To study worker tasks, I build on the occupational classification developed by Bernard et al. (2017); see Appendix A.6 for details.

I use two new and complementary micro data sources to measure robot adoption at the firm level. I first use a robot adoption firm survey conducted by Statistics Denmark in 2018. The survey asked a representative sample of Danish firms if they use industrial robots in production. Appendix A.1 provides details on the survey which had a response rate of 97 percent. Second, I leverage the fact that industrial robots are highly tradable goods to measure robot adoption from firm customs records. Almost all robots in the world are manufactured in Japan, South Korea, or Germany, and once such an imported robot crosses the country border, it is recorded by the customs authorities according to a 6-digit product code where one of the codes identifies "847950 Industrial Robots". Acemoglu and Restrepo (2018a) show that a country's imports of industrial robots correlate strongly with its total robot installments reported by the International Federation of Robotics (IFR). Appendix A.4 calculates that the share of imports in total robot investments in Denmark averaged 95 percent between 1993 and 2015. The Danish customs records are organized in the Foreign Trade Statistics Register (UHDI).

The main challenge in using the customs records is that a substantial share of machinery is imported through domestic distributors. In the case of industrial robots, there is an industry of robot integrators that specialize in importing robots and installing them at local production facilities. Appendix A.2 describes the robot supply chain and develops a procedure for identifying robot imports done by final adopters. I validate the sample selection procedure against the firm robot adoption survey in 2018 as well as a complete list of robot integrators and producers in Denmark. In total, I identify 454 robot adoption

events through direct imports.

The existing literature on industrial robots has mostly relied on an industry-level dataset compiled by the International Federation of Robotics (IFR) (Acemoglu and Restrepo, 2020; Dauth et al., 2018; Graetz and Michaels, 2018). Appendix A.3 shows that the micro data used in this paper align well with the IFR statistics both across industries and over time.

The customs records allow me to directly study what happens when firms adopt robots. However, when quantifying the aggregate effects of robots and for parts of the structural estimation, I also want to include the adoptions done through domestic distributors. Appendix A.5 describes how I supplement the customs records with the robot adoption survey (VITA) and the IFR statistics to measure robot adoptions that are sourced domestically.

The customs records (UHDI) and the robot adoption survey (VITA) use the same firm identifier as FirmStat and FIDA, allowing me to construct a matched employee-employer-robot dataset covering the Danish economy.

2.1 Stylized Facts on Firm Robot Adoption

In this section, I present two stylized facts that will inform the modeling choices in Section 3. The first fact concerns the observed lumpiness of firm robot expenditures, which motivates modeling robot adoption as a one-off decision. The second fact documents the non-random selection of firms into robot adoption, which informs the specification of a selection model for firm robot adoption.

Fact 1. Robot Adoption Is Lumpy

Table 1 reports summary statistics for the robot adoptions identified from firm customs records in Appendix A.2. The take-away from the table is that robot adoption is lumpy. Out of the sample adopters, 70.6 percent invest in a single year only, and the peak year of investment accounts on average for 90.7 percent of total firm robot expenditures. Adopting firms purchase robot machinery for an average of $311,000. This discrete nature of robot adoption motivates the choice in Section 3 to model robot adoption as a discrete choice problem.

Table 1: Firm Robot Investments

Adoptions (count)	454
Share of adopters with investments in one year only (percent)	70.6
Share of robot expenditures in max year (percent)	90.7
Robot machinery expenditures ($1,000)	311

Fact 2. Larger Firms Select into Robot Adoption

Table 2 shows firm outcomes for the robot adopters in the year prior to adoption. Column 2 ("Industry") reports average outcomes for non-adopters within the same two-digit industry-year cells as the robot adopters. Robot adopters are different from non-adopters along several dimensions, but the key feature that sets robot adopters apart is that they are substantially larger. The model in Section 3 rationalizes the selection into robot adoption by combining firm heterogeneity with fixed costs of adoption, such that it is the firms with the largest expected efficiency gains from industrial robots that will choose to adopt the technology.

Table 2: Firm Outcomes in the Year Before Robot Adoption

	Adopters	Industry	Matches	P-value A-M
log Sales	18.28 (0.07)	16.35 (0.07)	18.19 (0.07)	0.37
log Wage Bill	16.93 (0.07)	15.15 (0.07)	16.89 (0.07)	0.66
log Employment	4.06 (0.06)	2.4 (0.06)	4.02 (0.06)	0.66
Wage bill shares (percent)				
– Managers	12.5 (0.5)	9.1 (0.7)	11.0 (0.4)	0.02
– Tech	16.0 (0.9)	6.9 (0.6)	14.3 (0.8)	0.14
– Sales	12.2 (0.4)	10.5 (0.6)	12.5 (0.5)	0.64
– Support	7.5 (0.4)	4.9 (0.5)	7.8 (0.5)	0.69
– Transportation/warehousing	5.9 (0.5)	3.6 (0.5)	6.8 (0.5)	0.23
– Line workers (mostly production)	39.9 (1.1)	47 (1.4)	40.7 (1.0)	0.61
Joint orthogonality (F test)				0.25
Observations	454	454	454	908

Notes: "Joint orthogonality" represents a test of the joint hypothesis that all coefficients equal zero when the adopter indicator is regressed on the nine outcome variables in Table 2. Column 1 (Adopters) shows mean outcomes for robot adopters in the year before adoption. Column 2 (Industry) shows averages for randomly chosen non-adopters within the same industry-year cell as the adopters (one-to-one). Column 3 (Matches) shows averages for match firms within the same industry-year cell. These matches each have the minimum distance to an adopter with respect to log sales and production wage bill share (levels and two-year changes); see Appendix A.7.1 for details. Column 4 (P-value A-M) shows p-values for the null hypotheses that Adopters (column 1) and Matches (column 3) have the same population mean.

Once I match on firm sales and line worker wage bill shares in column 3 ("Matches"), the adopters look similar to the match firms on employment, wages, and wage bill shares across occupations. An F test of the joint hypothesis that none of the covariates in Table 2 predict robot adoption has a p-value of 0.25. Put differently, I cannot reject that robot adopters and matches indeed are observationally identical before adoption. The fact that adopters and match firms are balanced on these non-targeted outcomes provides sup-

portive evidence for a model assumption in Section 3 that robot adoption is driven by an adoption cost shock once selection based on observable firm heterogeneity is taken into account.

3 A Model of Firm Robot Adoption

In this section, I develop a partial-equilibrium model of a manufacturing firm's decision to adopt industrial robots. A firm in the model faces a dynamic choice of whether to adopt the robot technology and a sequence of static decisions to hire workers and use intermediate inputs for production. The optimal adoption decision trades off a sunk cost of robot adoption with gains in future profits from being able to operate the robot technology.

In Sections 3.1 and 3.2, I characterize the firm's static production problem taking the robot technology choice as given. In Section 3.3, I then characterize the firm's dynamic problem of adopting robot technology. The firm problem is linked to the worker's problem in general equilibrium but only through the path of wages. This separable structure allows me to study and estimate the firm model in isolation by conditioning on the observed path of wages, and postpone the specification of the worker's problem to Section 5.

3.1 Production Technology

A manufacturing firm j uses workers of different occupations $L \in \mathbb{R}_+^{|\mathcal{O}|}$ and intermediate inputs $M \in \mathbb{R}_+$ according to the CES production function

$$Y_{jt} = F(M_{jt}, L_{jt}|R_{jt}, \varphi_{jt}) = z_{Hjt} \left\{ M_{jt}^{\frac{\sigma-1}{\sigma}} + \sum_{o \in \mathcal{O}} z_{ojt}^{\frac{1}{\sigma}} L_{ojt}^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} \quad \text{with} \tag{1}$$

$$z_{Hjt} = \exp(\varphi_{Hjt} + \gamma_H R_{jt}) \tag{2}$$

$$z_{ojt} = \exp(\varphi_{ojt} + \gamma_o R_{jt}) \tag{3}$$

Firms are heterogeneous with respect to a vector of exogenous baseline productivities $\varphi \in \mathbb{R}^{\mathcal{O}+1}$ and an endogenous robot technology state $R \in \{0, 1\}$. The parameter γ_H captures the effect of robot technology on firm Hicks-neutral productivity z_H, and the parameters γ_o govern how robot technology changes the relative productivities of worker occupations in production z_o (measured relative to intermediate inputs M).[1]

In modeling robot adoption as a technology choice, I follow a growing literature arguing for task-based models to study automation (Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018b). Appendix B.1 derives the specification in Equation (1) from a micro-founded model in which robots substitute for production tasks performed by workers. I model robot technology as a binary state $R \in \{0, 1\}$ to reflect the fact that most robot users invest in robots in a single year only (Fact 1 from Section 2.1).

[1] Intermediate inputs M include all non-labor inputs including materials and conventional capital equipment. I measure payments to these intermediate inputs as the part of firm sales that are not paid to labor or profits. As Section 3.3 will make clear, industrial robots are different from other non-labor inputs in that their adoption involves a change of production technology that is subject to a sunk robot adoption cost.

3.2 Demand and Flow Profits

The firm faces an iso-elastic demand curve

$$Y_{jt} = Y_{Mt} \times (P_{jt}/P_{Mt})^{-\epsilon}, \tag{4}$$

where Y_{Mt} is the aggregate manufacturing demand and P_{Mt} is the manufacturing price index. The firms takes the vector of factor prices w_t as given, such that the flow profit function reads

$$\pi_t(R, \varphi) = \max_X \left\{ P_{Mt} Y_{Mt}^{\frac{1}{\epsilon}} F(X|R, \varphi)^{1-1/\epsilon} - w_t^T X \right\} = \Omega_t C_t(R, \varphi)^{1-\epsilon}, \tag{5}$$

where C_t denotes the unit cost function, Ω_t is a common profit shifter, and the static inputs are stacked into the vector $X = (M, L)$.[2] By lowering production costs C_t, the robot technology allows firms to scale up output and increase flow profits.

The key assumption in Equation (1) is that the production function admits a static factor demand system (satisfying Equation (5)) that is invertible in firm productivities. Invertibility allows me to control for unobserved firm productivities by matching on observed factor choices, similar to the proxy variable approach to production function estimation (Ackerberg et al., 2015; Levinsohn and Petrin, 2003; Olley and Pakes, 1996). Berry et al. (2013) show that a demand system is invertible if and only if it satisfies a "connected

[2]The unit cost function and profit shifter are given by the CES expressions

$$C_t(R, \varphi) = \frac{1}{z_H(R, \varphi)} \left\{ \sum_{x \in X} \left(w_{xt} / z_x(R, \varphi) \right)^{1-\sigma} \right\}^{\frac{1}{1-\sigma}}, \quad \Omega_{jt} = P_{Mt}^{\epsilon} Y_{Mt} (\epsilon - 1)^{(\epsilon-1)} \epsilon^{-\epsilon}. \tag{6}$$

substitutes" condition. The set of such production functions includes CES as in Equation (1), non-homothetic CES, nested CES, mixed CES, and translog. Appendix C.2.2 relaxes the robot technology effects in Equation (2)-(3) to a distributed lag model to account for any adjustment dynamics in the transition of firms to robot production. The demand curve in Equation (4) can be relaxed to an arbitrary downward-sloping function as considered in Doraszelski and Jaumandreu (2018). Appendix E derives an extension of the model where firms face upward-sloping labor supply curves and thus do not take wages as given in Equation (5).

3.3 Adoption of Robot Technology

The firm faces a dynamic decision about whether and when to adopt the robot technology R. The optimal adoption decision trades off a sunk cost of robot adoption with gains in future profits from being able to operate robot technology. The sunk adoption cost includes a common time-varying component c_t^R and an idiosyncratic component ε_{jt}^R. The adoption decision is essentially an optimal stopping problem that is reminiscent of the seminal work on bus engine replacement by Rust (1987). The value of a firm is represented by the Bellman equation

$$V_t(0,\varphi) = \max_{R\in\{0,1\}} \pi_t(0,\varphi) - (c_t^R + \varepsilon_{jt}^R) \times R + \beta \mathbb{E}_t V_{t+1}(R,\varphi') \tag{7}$$

$$V_t(1,\varphi) = \sum_{\tau=0}^{\infty} \beta^\tau \pi_{t+\tau}(1,\varphi_{t+\tau}). \tag{8}$$

Robot technology does not depreciate in the baseline specification of the model.[3]

[3] Appendix C.5 specifies and estimates a model extension in which robots deteriorate at a fixed rate.

Firm baseline productivities evolve according to the Markov process

$$\varphi_{jt+1} = g_t(\varphi_{jt}, ..., \varphi_{jt-k}) + \xi_{jt+1}, \quad \xi_{jt+1} \perp\!\!\!\perp (\varphi_{jt}, ..., \varphi_{jt-k}), (\varepsilon^R_{jt}, ..., \varepsilon^R_{jt-l}). \tag{9}$$

The idiosyncratic adoption cost shocks ε^R_{jt} are drawn i.i.d. from a cumulative distribution function F such that the probability that a firm adopts robot technology is

$$P_t(\Delta R_{jt+1} = 1) = F\left(\beta\left(\mathbb{E}_t V_{t+1}(1, \varphi_{jt+1}) - \mathbb{E}_t V_{t+1}(0, \varphi_{jt+1})\right) - c^R_t\right) \tag{10}$$

The multiplicative productivity effects of robots in Equations (2) and (3) imply that firms that operate on a larger scale will be better able to reap the benefits of robot technology. Combined with the fixed component of robot adoption costs c^R_t, this allows the model to rationalize the observed size premium in robot adoption (Fact 2 from Section 2.1). It is, however, worth noting that the model also allows for variable costs of robot adoption through the γ_o parameters. Robot production will, for example, be more intensive in intermediate inputs if γ_o is negative or require more tech workers if γ_T is positive. The adoption model also implies that larger firms will spend more on robots when they adopt because these firms will be willing to pay a higher idiosyncratic adoption cost ε^R_{jt}.

The robot adoption model features two key simplifying assumptions about robot investment behavior. First, robot adoption is treated as a one-off decision. This assumption is motivated by the observed lumpiness (Fact 1 in Section 2.1) whereby most robot users invest entirely in a single year. Appendix C.5 estimates a model extension in which robots deteriorate at a fixed rate, thereby leaving scope for replacement investments. Sec-

ond, firms cannot receive larger relative robot production effects γ by spending more on robots. The structural estimation in Section 4 will provide empirical evidence in support of this homogeneity assumption on the treatment effects of robot adoption.

Equation (7) entails a key timing assumption that robot adoption is decided one year in advance. Combined with the Markovian structure on the productivity process in Equation (9), this timing assumption will be key to separating out the causal impact of robot adoption on firm productivities in Section 4.[4]

4 Structural Estimation of Firm Robot Adoption

In this section, I estimate the robot adoption model presented in Section 3. The structure of the model allows me to estimate its parameters in sequence. In Sections 4.1 to 4.3, I estimate the parameters of firm production technologies without having to specify other parameters of the adoption model, including robot adoption costs. In Section 4.4, I then estimate the cost parameters of robot adoption. I set the elasticity of demand and the time discount factor to conventional values from the literature ($\epsilon = 4, \beta = 0.96$).[5]

[4]The timing assumption on investment decisions (a one-year time-to-build) combined with a Markov process for firm productivities is a common assumption in the production function estimation literature, including Olley and Pakes (1996) and Doraszelski and Jaumandreu (2013).

[5]I follow Bloom (2009) and Asker et al. (2014), who calibrate the elasticity of demand ϵ to 4 to reflect a markup on output prices of 1/3 and calibrate the annual discount rate β to the data reported in King and Rebelo (1999).

4.1 Elasticity of Substitution between Production Tasks

In this section, I estimate the elasticity of substitution between production tasks, σ. I distinguish between labor tasks of production workers, tech workers, and other workers.[6] To preview, I use the model structure to derive an instrumental variables strategy, and I estimate that tasks are complements in firm production.

The first-order conditions for cost minimization in Equation (5) imply that firm factor demands satisfy the following relationship

$$\log(L_{o'jt}) - \log(L_{ojt}) = -\sigma(\log(w_{o'jt}) - \log(w_{ojt})) + \log(z_{o'jt}) - \log(z_{ojt}) \tag{11}$$

The challenge in using Equation (11) to estimate σ is the classic simultaneity problem (Marschak and Andrews, 1944) that wages w_{jt} may be correlated with firm productivities z_{jt}, which constitute the regression error term in Equation (11). Appendix E derives a model in which firms face upward-sloping labor supply curves, thus creating an explicit link between firm productivities and wages.

I use the structure of the model in Section 3 to derive a rational expectations generalized method of moments (GMM) estimator that explicitly solves this simultaneity problem. The identification strategy builds on the insight of Doraszelski and Jaumandreu (2018) that the Markovian structure on firm productivities implies that past factor choices X_{jt-1} and prices w_{jt-1} must be uncorrelated with the current productivity innovations ξ_{jt}.

[6]The classification of worker tasks builds on the occupational grouping of Bernard et al. (2017); see Appendix A.6 for details.

This restriction allows me to estimate σ from the moment condition

$$\mathbb{E}\left[A_{oo'}(Q_{jt-1})(\xi_{ojt} - \xi_{o'jt})\right] = 0, \tag{12}$$

where $A_{oo'}$ is a vector function of the instruments Q_{jt-1}, including $\log(X_{jt-1})$ and $\log(w_{jt-1})$. The derivation of this moment condition closely follows Doraszelski and Jaumandreu (2018), and I therefore relegate the derivations to Appendix C.1. The key idea is to, first, break the productivity error term z_{jt} in Equation (11) into the predictable component g_{jt} and the innovation ξ_{jt}. Since firms behave with rational expectations, the unforeseeable innovations ξ_{jt} must be uncorrelated with past decisions and prices of firms. To the extent that lagged factor prices and decisions correlate with current factor prices, they thus constitute valid and relevant instruments for estimating the substitution elasiticity σ.

I estimate Equation (12) on the sample of non-adopters, which allows me to identify σ without specifying how robot technology affects firm productivities in Equations (2)-(3), which I separately estimate in Section 4.2. I estimate the moment conditions using a two-step GMM procedure with Appendix C.1 providing additional details on the estimation problem. The GMM estimate of the elasticity of task substitution σ is 0.49, which implies that tasks are complements in firm production. This estimate is based on the Danish matched employer-employee data from 1995 to 2015.

Table 3: Estimating the Elasticity of Substitution between Tasks in Production

	GMM
Elasticity of task substitution, σ	0.493
	(0.092)

To place this estimate in the literature, Doraszelski and Jaumandreu (2018) estimate

that the elasticity of substitution between labor and materials lies between 0.4 and 0.8, while Raval (2019) estimates that the elasticity of capital-labor substitution to falls between 0.3 and 0.5. There is, to my knowledge, no estimate in the existing literature of the micro elasticity of substitution between worker tasks, and one contribution of this section is to provide such an estimate.[7]

4.2 Robot Technology

In this section, I estimate the parameters of robot technology γ, a key input for evaluating the distributional impact of industrial robots. In Section 4.2.1, I first use the model in Section 3 to derive an identification strategy that is based on event studies of firm robot adoption. In Section 4.2.2, I then present the estimation results, which show that industrial robots increase production efficiency but cause a substantial bias in technology away from production workers and toward tech workers and intermediate inputs.

4.2.1 Identification of Robot Technology

This section describes my strategy for identifying the parameters of robot technology, γ. I first discuss the identification challenges that arise from the fact that firms endogenously select into robot adoption. I then use the adoption model developed in Section 3 to derive an identification strategy that deals with this selection problem.

First, from the invertibility of the factor demand system, I can recover firm productiv-

[7] An important reason for the absence of such an estimate is the lack of micro data on the labor tasks employed in firms. The detailed occupational codes in the Danish data are unusually rich in this regard.

ities from the first-order conditions to Equation (5)

$$z_{ojt} = l_{ojt} - m_{jt} + \sigma(\log w_{ojt} - \log w_{Mjt}) \tag{13}$$

$$z_{Hjt} = \frac{1}{\epsilon - 1} m_{jt} + \frac{\sigma}{\epsilon - 1} \log w_{Mjt} + \frac{(\sigma - \epsilon)}{(\sigma - 1)(\epsilon - 1)} \log \left\{ w_{Mjt}^{1-\sigma} + \sum_{o} z_{ojt} w_{ojt}^{1-\sigma} \right\} \tag{14}$$

where lower-case factor choices denote log transforms. With these productivities recovered, it is now tempting to use Equations (2)-(3) to run the regression

$$\log(z_{jt}) = \gamma R_{jt} + \varphi_{jt} \tag{15}$$

The issue with using Equation (15) as an estimating equation is that firms adopt robots R_{jt} based on their expected baseline productivities φ_{jt} (see Equation (22)), which exactly is the error term in Equation (15), thus creating selection bias. For example, simply comparing robot adopters to non-adopters in the cross-section will create bias because high baseline productivity firms are better able to overcome the fixed cost of robot adoption. Similarly, simply comparing a firm before and after robot adoption will be biased because firms tend to adopt robots when their baseline productivity is high or when they expect to face high demand for their products. Indeed, Fact 2 of Section 2.1 showed that robot adopters tend to be larger.

As I will show formally below, the dynamic adoption model of Section 3 gives me a way to confront this selection problem. The key idea is to match on observed firm factor choices leading up to adoption to control for selection into robot adoption based on heterogeneity in firm productivities. The reason why observably similar firms make

different decisions about robot adoption is then due to heterogeneity in the sunk costs of robot adoption ε_{jt}^{R}, which satisfies the exclusion restriction for identification in the model. The key identifying assumption is that observed factor choices are sufficient to control for firm productivities, and that there is no selection on unobservables that directly affect firm outcomes. The matching-based event study identification strategy reads as follows.

Identification Strategy (Parameters of Robot Technology γ)**.**

1. Take two firms with similar output and occupational wage bills in some initial k years.

2. In the following year, one of the firms adopts robots.

3. The differential paths of firm output and occupational wage bills identify the parameters of robot technology, γ .

The firm model in Section 3 falls into a general class of potential outcomes models for robot adoption. In these potential outcomes models, the assumptions for non-parametric identification of average treatment effects are well-understood (Imbens and Wooldridge, 2007). I first remind the reader of these general requirements for identification, and then show that they are satisfied in my adoption model. Finally, I show that the average treatment effects estimated by the event studies identify the robot technology model parameters of interest.

Note first that, since payments to intermediate inputs M are defined as the part of firm sales that is not paid to labor or profits (a constant markup on firm sales), matching on firm sales and occupational wage bills is equivalent in the model to matching on the full vector of firm factor bills, $X = (M, L)$.

In the model, a firm's factor demands $x_{jt} = (m_{jt}, l_{jt})$ can take two potential values, $(x_{jt}(0), x_{jt}(1))$, according to whether or not the firm has adopted robot technology. In the language of Rubin (1990), the two identifying assumptions are *unconfoundedness*

$$\left\{\Delta R_{jt+1} \perp\!\!\!\perp \left(x_{jt+1}(1), x_{jt+1}(0)\right)\right\} \mid \left(x_{jt-1}(0), .., x_{jt-k}(0)\right) \tag{A1}$$

and *overlap* in robot adoption

$$0 < P\left(\Delta R_{jt+1} = 1 \mid x_{jt-1}(0), .., x_{jt-k}(0)\right) < 1 \tag{A2}$$

Assumption (A1) requires that, once I condition on the path of factor choices that lead a firm to adopt robots in year t, the act of adoption must be independent of the firm's potential factor choice outcomes going forward. On top of this, Assumption (A2) requires that I can find another firm that experienced the same initial sequence of factor choices but did not adopt robots in year t. Under Assumptions (A1) and (A2), the difference in sample means between adopter and match firms identifies the average treatment effect of robot adoption (see Imbens and Wooldridge (2007))

$$\bar{x}^T_{t+1} - \bar{x}^C_{t+1} \xrightarrow{p} \mathbb{E}\left[x_{jt+1}(1) - x_{jt+1}(0) \mid j \in T\right], \tag{16}$$

where $\bar{x}^T$ and $\bar{x}^C$ denote the sample means for adopter and match firms, respectively.

Let us now see how the general identifying assumptions (A1) and (A2) derive from the adoption model in Section 3. First, by the invertibility of the factor demand system, I am implicitly conditioning on $(\varphi_{jt-1}, ..., \varphi_{jt-k})$ when I match on firm factor choices in

the k years that lead up to robot adoption (see Equations (13) and (14)).[8] Once I condition on $(\varphi_{jt-1}, ..., \varphi_{jt-k})$, firm future factor outcomes $(x_{jt+1}(0), x_{jt+1}(1))$ are driven solely by the productivity innovations ξ_{t+1} in Equation (9). Since these future productivity innovations are unforeseeable when firms choose to adopt robots in year t, the adoption model satisfies the *unconfoundedness* condition (A1).

Second, the probability of robot adoption in the model is given by

$$P_t(\Delta R_{jt} = 1 | \varphi_{jt-1}, ..., \varphi_{jt-k}) = F\left(\beta\left(\mathbb{E}_t V_{t+1}(1, \varphi_{jt+1}) - \mathbb{E}_t V_{t+1}(0, \varphi_{jt+1})\right) - c_t^R\right) \quad (17)$$

which lies strictly within the unit interval as long as the distribution of idiosyncratic adoption costs F has full support. The adoption model thus also satisfies the *overlap* condition (A2). Put into words, the identification strategy relies here on firm heterogeneity in the costs of robot adoption ε_{jt}^R driving otherwise similar firms to make different decisions about robot adoption.

Finally, from the model equations (2), (3), (13) and (14), we see that the treatment effects in Equation (16) identify the parameters of the robot technology

$$\gamma_o = z_{ojt}(1) - z_{ojt}(0) = \big(l_{ojt}(1) - l_{ojt}(0)\big) - \big(m_{jt}(1) - m_{jt}(0)\big) \quad (18)$$

$$\gamma_H = z_{ojt}(1) - z_{ojt}(0) \quad (19)$$

$$= \frac{1}{\epsilon - 1}\big(m_{jt}(1) - m_{jt}(0)\big) + \frac{(\sigma - \epsilon)}{(\sigma - 1)(\epsilon - 1)} \log\left\{\frac{w_{Mjt}^{1-\sigma} + \sum_o z_{ojt}(1) w_{ojt}^{1-\sigma}}{w_{Mjt}^{1-\sigma} + \sum_o z_{ojt}(0) w_{ojt}^{1-\sigma}}\right\} \quad (20)$$

[8] If wages are firm-specific, the identification strategy also requires me to match on wages. In the analysis below, I match on factor choices, and then show that the firms also match on wages. The non-targeted match on wages is as an overidentification check of the model assumption that robot adopters do not pay wage premiums.

The identification of γ_H requires the values of the factor augmenting productivities z_{ojt} which at this point can be readily recovered from Equation (13).

4.2.2 Estimation Results

The identification strategy presented above suggests matching robot adopters to comparison firms with a similar path of factor choices leading up to the adoption event. The match firms found in column 3 of Table 2 in Section 2.1 satisfy exactly these criteria. To recap, I found these firms by matching each robot adopter to a non-adopter firm that operated in the same two-digit industry and had a similar trajectory of firm sales and line worker wage bill shares in the three years that led up to adoption.[9] I then showed that these firms were similar to the robot adopters on the full vector of factor choices as required by the identification strategy above.[10]

Once I have matched firms based on their factor choices leading up to robot adoption, the model in Section 3 implies that the act of adoption is driven by the idiosyncratic cost shock ε_{jt}^R that is independent of all other drivers of firm outcomes. The fact that the adopter and match firms are similar on several non-targeted outcomes in Table 2 provides evidence in support of this identifying assumption. The fact that the firms pay similar wages, in particular, provides an overidentification check of the model assumption that robot adopters do not pay wage premiums.

To ease the exposition, I presented the adoption model in Section 3 assuming that the

[9] I use an Exact-Mahalabonis matching procedure described in Appendix A.7.1. The three-year match window allows for firm productivities in Equation (9) to follow an arbitrary Markov chain of length three.

[10] A test of the joint hypothesis that none of the covariates in Table 2 (firm sales, employment, and occupational wage bills) predict robot adoption has a p-value of 0.25 (reported in the second-to-last row of the table).

productivity effects of robotization γ manifest fully within the first year of adoption; see Equations (2)-(3). When taking the model to the data, I allow for the possibility that firms take a longer time to fully adjust to robot production. In practice, I track firm outcomes for four years after robot adoption. This, however, opens the possibility that some of the control firms may have also adopted robots in the post-event time window. Appendix Figure C.1 shows that around 10 percent of control firms adopted robots four years after the event year, which works against finding an effect of robot adoption in the reduced form of the event studies. I take this change in treatment status into account when estimating the model parameters.[11]

Figures 1 and 2 show the main results from the estimation of robot technology. The figures display the differential paths of firm size and factor choices around robot adoption as prescribed by the identification strategy above. The blue lines represent raw data and the dashed orange lines show the model fit.[12] As I showed in Section 4.2.1, these reduced-form effects exactly identify the parameters of robot technology γ.

I estimate the parameters of robot technology to match the reduced-form moments four years after robot adoption. I choose the four-year horizon to account for the smoother transition path to robot production found in the data. This transition path likely reflects complementary investments that occur post adoption but that the model assumes are incurred immediately upon adoption. Appendix C.2.2 generalizes the model in Section 3 to account for these dynamic adjustments to robot production by allowing the productiv-

[11]The model-implied correction is the Wald estimator used in the treatment effects literature to convert intention-to-treat (ITT) effects into treatment-on-the-treated (TOT) estimates; see Angrist and Pischke (2008).

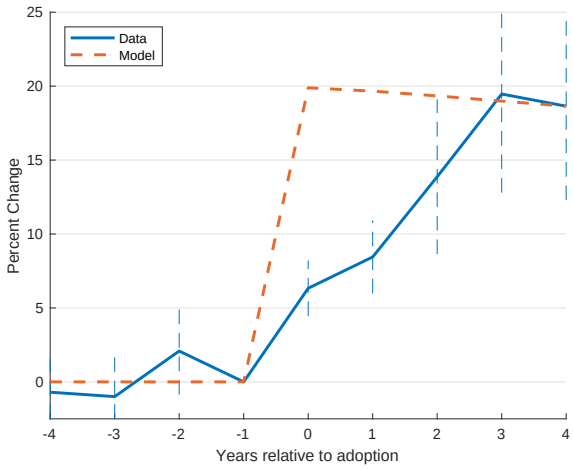
[12]Appendix C.2.1 describes the econometric specification that generates the point estimates and confidence intervals plotted in Figures 1 and 2.

ity effects of robot adoption in Equations (2)-(3) to follow a distributed lag model. The appendix section estimates the full dynamic path of robot productivity effects. This generalization adds to the computational complexity of the model by requiring me to keep track of the years since robot adoption when solving the firm's dynamic programming problem. With the aim of keeping the firm's state space tractable when solving the general equilibrium model in Section 6, I abstract from these dynamic adjustment processes and instead match directly on the reduced-form effects four years after robot adoption.

The figures show that the model-simulated diff-in-diffs tend to drift back toward zero in the years following adoption. This post-event drift toward zero reflects the control firms that adopt robots in the post-event time window (orange line in Appendix Figure C.1).

Figure 1: Firm Outcomes Around Robot Adoption (Matching Diff-in-Diff)

(a) Sales



(b) Wage Bill

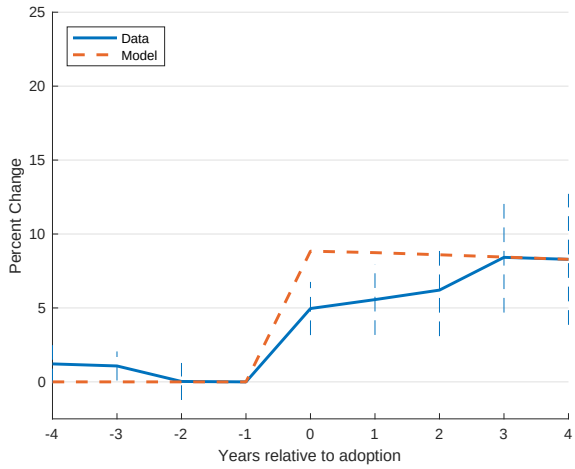


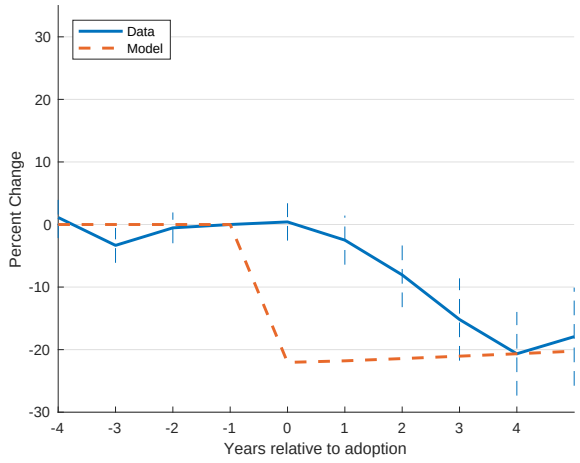
Figure 1(a) shows that the average firm's sales increase 20 percent around robot adoption. Through the lens of the structural model, this sales effect implies that robot technology increases firm production efficiency by around 7 percent, given the calibrated elastic-

ity of firm demand ϵ. Figure 1(b) shows that the wage bill increases by 8 percent around robot adoption. The wage bill increase is less than the 20 percent sales effect in Panel (a), and implies that the substitution effects of robot adoption on labor γ_o on average are negative.

Figure 2 decomposes the wage bill effects in Figure 1(b) by occupations. Production workers include tasks from welding to assembly, while tech workers include engineers, researchers, and skilled technicians. Panel (a) of Figure 2 shows that the demand for production workers falls by around 20 percent around robot adoption, while Panel (b) shows that the demand for tech workers simultaneously increases by around 30 percent. This shift of labor demand away from the production line and toward the tech department implies that robot adoption lowers the relative productivity of production workers ($\hat{\gamma}_P = -0.461$) but increases the relative productivity of tech workers ($\hat{\gamma}_T = 0.043$).

Figure 2: Firm Wage Bills Around Robot Adoption (Matching Diff-in-Diff)

(a) Production Workers



(b) Tech Workers

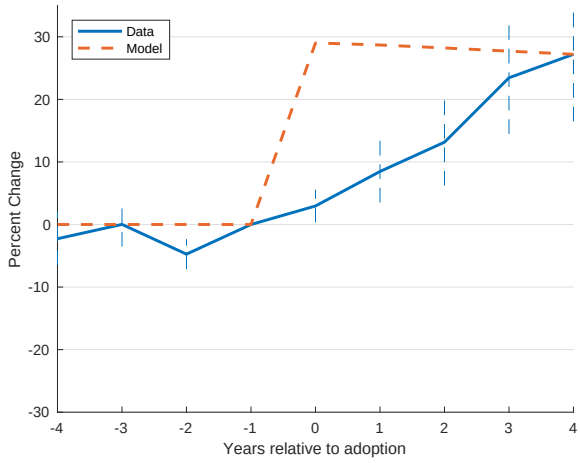


Table 4 summarizes the estimated parameters of robot technology.

Table 4: Estimated Parameters of Robot Technology

Parameter	Description	Estimated Value
γ_P	Production worker augmenting robot productivity	-0.461
γ_T	Tech worker augmenting robot productivity	0.043
γ_O	Other worker augmenting robot productivity	-0.115
γ_H	Hicks-neutral robot productivity (normalized)	0.066

Notes: The relative productivity effects γ_o are measured relative to intermediate inputs. The parameter γ_H is normalized such that a zero sales effect of robot adoption would imply a value γ_H of zero.

The reduced-form effects in Figure 1 align well with Koch et al. (2019), who find that robot adoption increases output 20-25 percent and lowers labor costs per unit produced among Spanish manufacturing firms. It is worth keeping in mind that the reduced-form effects in Figures 1 and 2 only identify the partial effects of one firm adopting industrial robots, and that any general equilibrium effects of robotization are differenced out in the figures. The general equilibrium model in Section 6 will fit these partial effects but also take into account general equilibrium interactions in product and labor markets to be able to quantify what happens when many firms in the economy adopt industrial robots.

4.3 Baseline Technology

Baseline productivities φ_{jt} are structural residuals that capture changes in firm production technology that are not due to robot adoption. I can now recover these baseline productivities by inverting the model equations. To be precise, with the robot technology parameters γ estimated in Section 4.2.2 and firm productivities z_{jt} recovered from Equations (13) and (14), I can use Equations (2) and (3) to retrieve baseline productivities φ_{jt}.

To solve their forward-looking problem of robot adoption, firms must form expectations about their future productivities. To estimate this robot adoption problem, I specify that firm productivities (Equation (9)) follow a first-order vector autoregression VAR(1) with Gaussian innovations.

$$\varphi_{jt} = \mu_t + \Pi\varphi_{jt-1} + \xi_{jt}, \quad \text{with} \quad \xi_{jt} \overset{iid}{\sim} \mathcal{N}(0, \Sigma). \tag{21}$$

The unknown parameters (μ_t, Π, Σ) in Equation (21) can readily be estimated using either maximum likelihood or three-stage least squares.

The general equilibrium model in Section 6 restricts the labor-augmenting part of baseline productivities to a common time-varying parameter vector φ_{ot}. This simplification is done to keep the firm's state space tractable and to home in on the key size dimension that sets robot adopters apart from non-adopters (Fact 2 of Section 2.1).[13] Appendix C.3.1 calibrates the path of common labor-augmenting productivities to match the path of manufacturing factor shares taking into account the diffusion of industrial robots. Appendix C.3.2 reports the results from estimating the productivity process in Equation (21). When solving the dynamic programming problem of robot adoption, I discretize the estimated baseline productivity process using the Tauchen (1986) method.

[13]The size premium in robot adoption is rationalized by the Hicks-neutral component of firm heterogeneity φ_{Hjt} which is left unrestricted. To be clear, the homogeneity restriction on firm baseline labor-augmenting productivities φ_{ot} is imposed solely for computational tractability: it does not alter the preceding analysis and can be relaxed without causing any conceptual or data complications.

4.4 Robot Adoption Costs

In this section, I estimate the costs of robot adoption. I first parameterize the path of common costs c_t^R and the distribution of idiosyncratic costs F, and then estimate their parameters to match the empirical robot diffusion curve and the observed firm size premium in robot adoption. To preview, I find that the model is able to generate the empirical S-shape in robot diffusion over time as well as the observed size premium of robot adopters, and that the estimated adoption costs align well with external cost measures.

I specify the idiosyncratic adoption cost shocks ε_{jt}^R to be drawn from a logistic distribution $F \sim \text{Logistic}(0, \nu)$ such that the probability that a firm adopts robot technology (Equation (10)) takes the form

$$P_t(\Delta R_{jt+1} = 1) = \frac{\exp(\frac{1}{\nu}(-c_t^R + \beta \mathbb{I}\mathbb{E}_t V_{t+1}(1, \varphi_{jt+1})))}{\exp(\frac{1}{\nu}(-c_t^R + \beta \mathbb{I}\mathbb{E}_t V_{t+1}(1, \varphi_{jt+1}))) + \exp(\frac{1}{\nu}\beta \mathbb{I}\mathbb{E}_t V_{t+1}(0, \varphi_{jt+1}))} \tag{22}$$

To develop intuition for the estimation strategy that I adopt here, note that Equation (22) implies a linear relationship between the log odds ratio of robot adoption and the expected gain in future profits from operating industrial robots.

$$\log\left(\frac{P_t(\Delta R_{jt+1} = 1)}{1 - P_t(\Delta R_{jt+1} = 1)}\right) = -\frac{c_t^R}{\nu} + \frac{1}{\nu} \times \left(\beta \mathbb{I}\mathbb{E} V_{t+1}(1, \varphi_{jt+1}) - \beta \mathbb{I}\mathbb{E}_t V_{t+1}(0, \varphi_{jt+1})\right) \tag{23}$$

Equation (23) shows that the common cost c_t^R governs the rate of robot diffusion, while the sensitivity of robot adoption to future profit gains is inversely linked to the dispersion

parameter ν.[14] Since larger firms are the ones that can better scale up production to reap the benefits of robot technology, and thus enjoy larger profit gains when adopting robots, it follows that the size premium in robot adoption is also inversely tied to ν. Following on this intuition, I develop a simulation-based estimator that entails searching for the adoption cost parameters, c_t^R and ν, that bring the model as close as possible to the observed robot diffusion curve and size premium in robot adoption.

I structure the exposition in two steps. In Section 4.4.1, I estimate the path of common adoption costs c_t^R to match the empirical robot diffusion curve, conditional on an estimate of ν. In Section 4.4.2, I then estimate the dispersion parameter ν to match the observed size premium in robot adoption. The final estimation procedure stacks the moments and estimates the parameters simultaneously using the method of simulated moments (MSM). Appendix C.4 provides details on the MSM estimation procedure.

4.4.1 Common Adoption Costs over Time

I estimate the path of common adoption costs $\{c_t^R\}_{t=0}^T$ to bring the model as close as possible to the observed robot diffusion curve. In particular, I parameterize the adoption cost schedule to be log-linear in time,

$$c_t^R = \exp(c_0^R + c_1^R \times t), \tag{25}$$

[14] By inverting continuation values from choice probabilities as in Arcidiacono and Miller (2011), I can rewrite Equation (23) as follows

$$\beta \log P_{t+1} - \log \frac{P_t}{1-P_t} = \frac{1}{\nu}(\beta c_{t+1}^R - c_t^R) - \frac{1}{\nu}\beta(\pi_{t+1}(1,\varphi') - \pi_{t+1}(0,\varphi')) \tag{24}$$

Equation (24) clarifies that the acceleration in robot diffusion pins down the change in robot adoption costs c_t^R over time, while $\frac{1}{\nu}$ measures the sensitivity of adoption to future profit flows.

and then search over a grid of intercepts c_0^R and slopes c_1^R to minimize the distance between the simulated and empirical diffusion curve. That is, for each pair of intercept and slope (c_0^R, c_1^R), I solve the dynamic programming problem of the firm, simulate the economy, and calculate the in-sample deviation to the empirical diffusion curve. The MSM estimator is the intercept-slope pair that brings the simulated diffusion curve the closest to the data. Appendix F.1 describes formally how to solve the dynamic programming problem of the firm. Put briefly, I first set a time horizon T sufficiently far in the future, such that robots are fully diffused by then. I then start at T, and solve the stationary, infinite horizon dynamic programming problem by iterating on the Bellman equation. I then solve for continuation values in $T-1, T-2, ...$, back to the first period using backward induction. With the continuation values in hand, I can simulate firms forward using the adoption policy functions, and verify that industrial robots have actually diffused fully by time T.

Figure 3(a) compares the fit of the estimated adoption curve, and Figure 3(b) plots the MSM estimate for the path of adoption costs. The common component of robot adoption costs amounts to 0.9 times the adopter firms' sales in 2019. This is, however, not the average sunk cost $c_t^R + \varepsilon_{jt}^R$ borne by adopters because firms select into robot adoption based on their idiosyncratic adoption cost ε_{jt}^R. Conditional on adoption, the total adoption cost amounts to around 10 percent of adopter sales.[15] These are the costs needed to rationalize the fact that, despite enjoying substantial sales gains upon robotization, only 31 percent

[15]Following Dubin and McFadden (1984), the expected cost borne by adopting firms may be calculated as

$$E(c_t^R + \varepsilon_{jt}^R | \Delta R_{jt+1} = 1) = c_t^R + \nu \left(\log P_t(\Delta R_{t+1} = 1) + \frac{P_t(\Delta R_{t+1} = 1)}{1 - P_t(\Delta R_{t+1} = 1)} \log P_t(\Delta R_{t+1} = 1) \right)$$

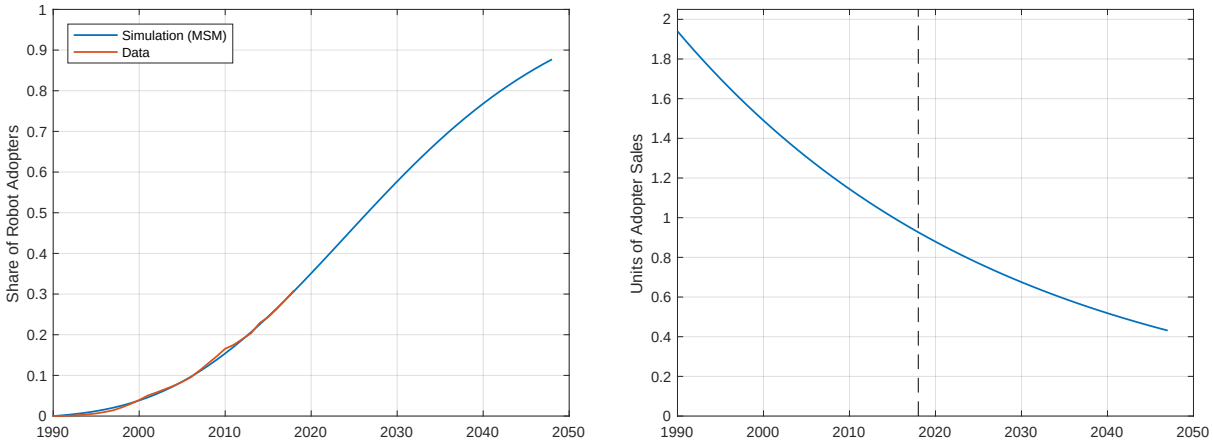
of manufacturing firms have adopted industrial robots almost 30 years after their arrival.

Figure 3: Estimating Adoption Costs on the Robot Diffusion Curve

(a) Robot Diffusion Curve

(b) MSM Estimate of Adoption Costs

$\hat{c}_t^R = \exp(\hat{c}_0^R + \hat{c}_1^R \times t)$



Note: Firm sales (the units in Panel (b)) are an average of adopter sales measured over the full simulation period.

One notable feature of Figure 3 is that, despite the log-linear schedule for adoption costs, the model (blue line in Panel (a)) is able to generate the S-shaped diffusion curve commonly found in the literature on technology adoption (Griliches, 1957). This can be seen as an overidentification check of the estimated adoption model. The model-simulated S-shape reflects the combination of a Bell-shaped distribution for firm productivity and a model where robot adoption is driven by threshold crossing in firm productivity. The Gaussian cumulative distribution function for baseline Hicks-neutral productivity φ_H naturally gives rise to a tail of technology leaders, a bigger mass of followers, and a tail of laggards, as implied by an S-shaped diffusion curve.

The MSM adoption cost estimate is an inferred cost that not only includes the monetary price of the robot machine but also expenditures for installation, the hassle of robot

adoption and production reorganization, as well as changing accessibility of industrial robots. Still, we may ask how the inferred adoption cost from my estimation procedure compares to external measures of robot investment costs. Table 1 showed that robot adopters on average spend a total of $311,000 on robot machinery. A rule of thumb is that machinery expenditures account for a third of the total cost of a robotic system that also includes expenditures for installation and integration (International Federation of Robotics, 2018). Taken together, this suggests that the monetary cost of robot adoption falls around $1 million. This number is slightly smaller than, but in the ballpark of, the inferred cost for adopters ($c_t^R + \varepsilon_{jt}^R$) of around 10 percent of firm sales. Appendix C.4.3 shows further that the estimated rate of change in adoption costs c_t^R aligns well with the robot machine expenditures reported on customs records of adopting firms.

Importantly, the MSM estimation procedure also identifies the path of future adoption costs that are consistent with the observed adoption behavior. This future path of adoption costs will be key to evaluating the effects of imposing a robot tax in Section 6.3.

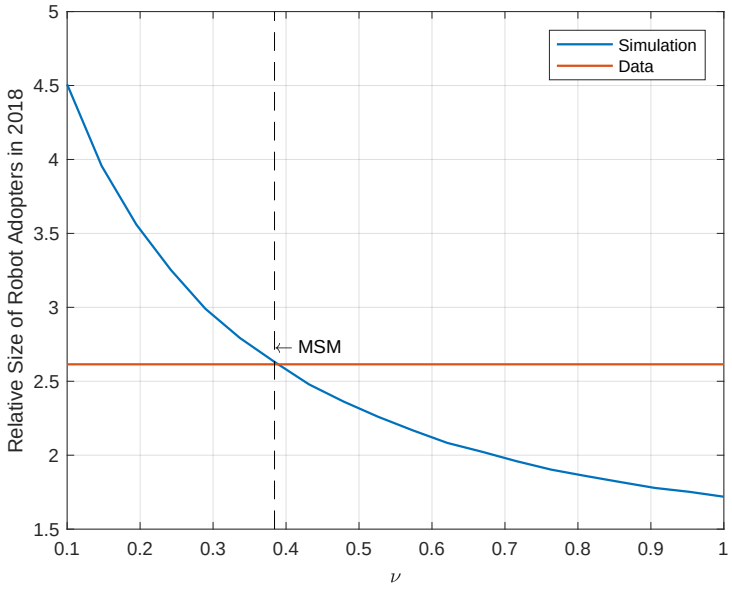
4.4.2 Variance of Idiosyncratic Adoption Costs

I estimate the dispersion in idiosyncratic adoption costs ν to match the observed size premium in robot adoption. Robot adopters were on average 2.61 times larger than non-adopter firms in 2018. The MSM procedure estimates ν to be 0.384, which delivers a simulated size premium of 2.61 in 2018. Figure 4 shows how the adopter size premium moment pins down the parameter ν by plotting the simulated size premium for varying values of ν.

To put this size premium into perspective, had selection into robot adoption been

unrelated to firm size ($\nu \to \infty$), the adopter premium would only have reflected the 20 percent sales effect estimated in Section 4.2. At the other extreme, without heterogeneity in adoption costs ($\nu \to 0$), robot adopters would have been around 6 times larger than non-adopters in 2018.[16] These estimates suggest that, while there is clear selection into robot adoption based on firm size (Fact 2 of Section 2.1), there is still ample heterogeneity in adoption costs ε_{jt}^{R}, leading observationally similar firms to make different decisions about robot adoption.

Figure 4: Size Premium of Robot Adopters for Varying Adoption Cost Dispersion ν



[16]The sales share of robot adopters in manufacturing was 53.9 percent in the data and in the model in 2018. In comparison, if firms did not select into robot adoption based on their size ($\nu \to \infty$) then the sales share of robot adopters would have been 34.5 percent. At the other extreme, without heterogeneity in adoption costs ($\nu \to 0$), the sales share would have been 72.8 percent.

5 The Labor Supply Block

This section presents the labor supply block of the general equilibrium model. I incorporate this labor supply module into the general equilibrium model in Section 6 to allow for a labor supply response to industrial robots where workers move out of adversely affected occupations. I use here a dynamic occupational choice model that represents the state-of-the-art for studying labor market dynamics in response to trade liberalizations (Dix-Carneiro, 2014; McLaren, 2017; Traiberman, 2019).

A key property of the general equilibrium is that the worker and firm problems are separable conditional on the path of wages. This block separable structure allows me to study and estimate the labor supply model now without reconsidering the firm's problem from Section 3 by conditioning on the observed path of wages.

The labor force consists of overlapping generations of heterogeneous workers as in Lee and Wolpin (2006). Workers enter the labor market at age 25 with an educational skill level $s \in \{\text{Low}, \text{Mid}, \text{High}\}$ and retire at age 65. In each year before retirement, workers face a choice of which occupation o to work in. This labor supply decision is dynamic in two ways. First, it is costly for workers to switch occupations. Second, workers may accumulate occupation-specific human capital on the job that is not transferable to other occupations. I allow labor markets to be segmented by occupation (production, tech, and other) and sector of employment (manufacturing and services).

A worker i of age a in occupation o in year t earns the product of a competitive occupational skill price, w_{ot}, and her human capital, H_{oit}. Her occupational human capital is

given by

$$\log(H_{oit}) = \beta_s^o s_{it} + \beta_1^o a_{it} + \beta_2^o a_{it}^2 + \beta_3^o \text{ten}_{oit} + \varsigma_{it} \tag{26}$$

where ten_o denotes tenure in occupation o, and $\varsigma_{it} \overset{iid}{\sim} \mathcal{N}(0, \sigma_h^2)$ is an ex-post productivity shock.

The worker's choice of occupation is an investment decision that trades off a sunk cost of switching occupations with future gains in wages and amenities of being employed in a new occupation. The occupational choice problem is represented by the Bellman equation

$$v_t(o, s, a, \text{ten}) = \max_{o' \in \mathcal{O}} \quad \log(w_{ot} H_o(s, a, \text{ten})) + \eta_{ot} - (c_{oo'}(s, a) + \varepsilon_{o'}) \tag{27}$$

$$+ \mathbf{1}_{\{a<65\}} \beta \mathbb{E}_t v_{t+1} \left(o', s, a+1, \mathbf{1}_{\{o'=o\}} (\text{ten}+1) \right) \tag{28}$$

where η_{ot} is a non-monetary amenity of working in occupation o, and $\varepsilon_o \overset{iid}{\sim} \text{GEV1}(\rho)$ is an idiosyncratic occupational switching cost shock. Income is implicitly assumed to be fully consumed in each period, and workers receive logarithmic flow utility of consumption. The occupational switching cost depends on the bilateral pair of current and prospective occupations, as well as the worker's age and skill

$$c_{oo'}(s, a) = c_{oo'} \exp \left\{ \alpha_s s + \alpha_1 \times a + \alpha_2 \times a^2 \right\} \tag{29}$$

I stack the worker state variables into the vector $\omega = (s, a, \text{ten}, o)'$.

5.1 Estimation of Labor Supply Parameters

I structurally estimate the labor supply model in Equations (26)-(28) using administrative data on the career paths of Danish workers. My approach to measurement and estimation follows closely Traiberman (2019). I describe the estimation procedures below, and relegate the data description and estimation results to Appendices D.1 and D.2. To preview, the estimate show that production workers face steep barriers to switching into tech occupations, that it is easier for workers to switch sectors instead of occupations, that workers accumulate specific human capital on the job that is not transferable to other occupations, and that older workers find it more costly to reallocate in the labor market.

5.1.1 Human Capital Function

I estimate the human capital function in Equation (26) using a Mincer regression of log earnings on worker skill, age, and occupational tenure.

$$\log(\text{Earnings}_{it}) = \log(w_{ot}) + \beta_s^o s_{it} + \beta_1^o a_{it} + \beta_2^o a_{it}^2 + \beta_3^o \text{ten}_{oit} + \varsigma_{it}, \tag{30}$$

where Earnings_{it} denotes labor earnings of worker i in year t, and w_{ot} is an occupation-time fixed effect. The key model assumption that enables me to identify the human capital parameters β in this regression is that workers cannot select on the productivity shock ς when choosing occupation or education. Appendix Table D.1 provides the OLS estimation results, which align with estimates from the existing literature (Ashournia, 2017; Dix-Carneiro, 2014; Traiberman, 2019).

5.1.2 Occupational Switching Costs

I estimate the occupational switching costs $c_{oo'}$ on observed worker transition and a conditional choice probability (CCP) estimator adapted from Traiberman (2019). The estimator exploits the finite dependence in the labor supply model to difference out unobserved continuation values by comparing workers who start and end in the same states (Arcidiacono and Miller, 2011).

The occupational choice model in Equation (27) implies that the difference in the (discounted) probabilities of observing a worker in occupation o first switching into occupation o' and then transitioning into occupation o'' compared to observing the worker first staying in occupation o and then transitioning into occupation o'' is

$$\log \frac{\pi_t(oo'|\omega)}{\pi_t(oo|\omega)} + \beta \log \frac{\pi_{t+1}(o'o''|\omega')}{\pi_{t+1}(oo''|\omega'')} = -\frac{1}{\rho} c_{oo'}(\omega) - \frac{\beta}{\rho}(c_{o'o''}(\omega') - c_{oo''}(\omega'')) \tag{31}$$

$$+ \frac{\beta}{\rho}\left(\log(w_{o't+1} H_{o'}(\omega')) - \log(w_{ot+1} H_o(\omega''))\right) \tag{32}$$

$$+ \frac{\beta}{\rho}(\eta_{o'} - \eta_o) + \zeta_{oo'o''t} \tag{33}$$

where $\pi_t(oo'|\omega)$ is the transition rate from occupation o to o' of workers with characteristics ω, H_o and w_{ot} are the human capital function and occupational skill prices estimated in Equation (30), and ξ is a mean-zero expectational error that is uncorrelated with the remaining RHS variables.

The occupational switching costs $c_{oo'}$ are identified off the excess likelihood of observing a worker staying in his own occupation from one year to the other, once his expected earnings differentials across occupations are controlled for. The occupational preference

shock variance ρ is estimated as the inverse elasticity of occupational switching with respect to expected earnings differentials.

The key model assumption in Equations (31)-(33) is that occupational switching is a renewal action that clears past choices from a worker's state. Combining this assumption with the Hotz-Miller inversion of continuation values from choice probabilities (Hotz and Miller, 1993) allows me to cancel out continuation values.[17]

Equations (31)-(33) constitute a system of non-linear regressions that identify the switching cost function $c_{oo'}$ and the preference shock variance ρ. Appendix D.2.1 describes the computational implementation of the estimation procedure. Appendix Tables D.2 and D.3 present the non-linear least squares (NLLS) estimation results. The estimates show that production workers face steep barriers to switching into tech occupations, that workers find it easier to switch sector within the same occupation, and that older workers find it more costly to reallocate in the labor market. The estimated switching cost magnitudes are in the range of those found in the existing literature.

The NLLS procedure tightly estimate all the occupational choice parameters, except for the preference shock variance ρ. In the current setup, the estimate of ρ greatly exceeds estimates in the existing literature. Since the labor supply responses to industrial robots are inversely related to this dispersion parameter, I choose to instead use a central estimate in the literature of ρ equal to 2. This value falls in between the estimates in Dix-Carneiro (2014), Ashournia (2017), Artuç et al. (2010), Caliendo et al. (2019), and Traiberman (2019).

[17]The derivation of Equations (31)-(33) closely follows Traiberman (2019), who estimates a richer model of labor supply that also accounts for unobserved (to the econometrician) types of workers.

5.1.3 Occupational Amenities

I estimate the path of occupational amenities η_{ot} to match the time series of employment shares across occupations. Appendix D.2.2 provides details on this estimation step.

6 Counterfactual Experiments

This section conducts counterfactual experiments to assess the general equilibrium impacts of industrial robots. I first present a general equilibrium model that unites the firm model from Section 3 with the worker model from Section 5. Section 6.1 defines the general equilibrium and develops a fixed-point algorithm for solving the equilibrium that features two-sided heterogeneity and dynamics. Section 6.2 uses the general equilibrium model to quantify how the arrival of industrial robots has affected the distribution of worker welfare. Section 6.3 evaluates the dynamic incidence of a robot tax.

6.1 Closing the General Equilibrium Model

The economy consists of a manufacturing sector and a service sector. The manufacturing sector consists of a mass $\mu_t^F(R, \varphi)$ of firms that are monopolistically competitive in product markets, pricetakers in factor markets, and otherwise operate as specified in Section 3.[18] Services are produced with a Cobb-Douglas technology and supplied competitively,

$$Y_{st} = z_{st} M_{st}^{\alpha_M^s} \prod_{o \in \mathcal{O}} L_{ost}^{\alpha_o^s} \tag{34}$$

[18]The baseline mass of firms $\mu_t^F(\cdot, \varphi)$ is taken as given but its distribution over the robot technology state R evolves endogenously according to the equilibrium robot adoption model.

The economy is populated by a mass $\mu_t^W(\omega)$ of workers who supply labor as specified in Section 5, and consume the final output bundle

$$Y_t = Y_{Mt}^{\mu} Y_{St}^{1-\mu} \quad \text{with} \quad Y_{Mt} = \left[\int Y(R,\varphi)^{\frac{\epsilon-1}{\epsilon}} d\mu_t^F(R,\varphi) \right]^{\frac{\epsilon}{\epsilon-1}} \tag{35}$$

I model Denmark, a country of less than 6 million people located in the European free trade zone, as a small open economy. Intermediate inputs M are imported at world price w_{Mt}, which the Danish economy takes as given, and trade is balanced. The robot adoption cost c_t^R is determined on the world market for industrial robots and is thus exogenous to local conditions in Denmark. The general equilibrium of the economy is defined as follows.

Definition 1 (Dynamic General Equilibrium)**.** A dynamic general equilibrium of the economy is a path of factor prices $\{w_t\}_t$, distributions of firm and worker states $\{\mu_t^F(R,\varphi), \mu_t^W(\omega)\}_t$, and policy functions $\{R_t(0,\varphi)\}_t$, $\{o_t'(\omega)\}_t$, such that taking the schedule of adoption costs $\{c_t^R\}_t$ and the price of intermediate inputs $\{w_{Mt}\}_t$ as given

1. Firms adopt robots to maximize expected discounted profits (Equation (7)) and demand static inputs to maximize profits period-by-period (Equation (5)).
2. Workers choose occupations to maximize expected present values (Equation (27)).
3. Labor markets clear (segmented by occupations and sectors)

$$\int L_{ot}(R,\varphi) d\mu_t^F(R,\varphi) = \int_\omega H_o(\omega) d\mu_t^W(\omega|M) \tag{36}$$

$$L_{ost} = \int_\omega H_o(\omega) d\mu_t^W(\omega|S), \tag{37}$$

where $L_{ot}(R, \varphi)$ is the static labor demand function satisfying Equation (5).

4. Firm output markets clear and trade is balanced.

$$Y_t = C_t + w_M M_t \tag{38}$$

where $M_t = \int M_t(R, \varphi) d\mu_t^F(R, \varphi) + M_{st}$ and $C_t = \sum_o w_{ot} L_{ot}^S + \Pi_t$. Equation (38) states that expenditures on intermediate input imports equal revenues from final goods exports.

5. The evolution of the distributions of firm and worker states $\{\mu_t^F, \mu_t^W\}_t$ is consistent with the policy functions $\{R_t(0, \varphi), o_t'(\omega)\}_t$.

A key property of the general equilibrium is that the firm and worker programs are separable conditional on the path of wages. This block separability breaks the curse of dimensionality where firm variables become states for the worker, and worker variables become states for the firm. The myriad of individual decisions taken by heterogeneous firms and workers is instead summarized into one aggregate state vector – the path of wages – which agents have perfect foresight about, up to unanticipated aggregate shocks to the economy. The block separable structure enables me to incorporate the rich firm and worker heterogeneity estimated in Sections 4 and 5, and still be able to compute the dynamic general equilibrium. In particular, the estimated general equilibrium model will fit the partial effects of firm robot adoption identified in Section 4 but also take into account how robotization affects non-adopter firms through product and labor markets, as well as the ability of workers to switch out of adversely impacted occupations.

I solve for the transitional dynamics of the economy where baseline productivities $\{\varphi_{jt}, z_{st}\}$, amenities $\{\eta_{ot}\}$, and robot adoption costs $\{c_t^R\}$ all have t-subscripts and are the time-varying fundamentals driving the system over time. The baseline estimated model perfectly matches the path of manufacturing factor bills (Appendix Figure C.3) and occupational employment shares (Appendix Figure C.3) observed in Denmark over time. I calibrate μ to match the manufacturing share in total output of the Danish economy and α_s to match the evolution of factor cost shares outside of manufacturing. Appendix Table G.1 provides a summary of the parameters of the general equilibrium model, as well as the moments used to estimate their values.

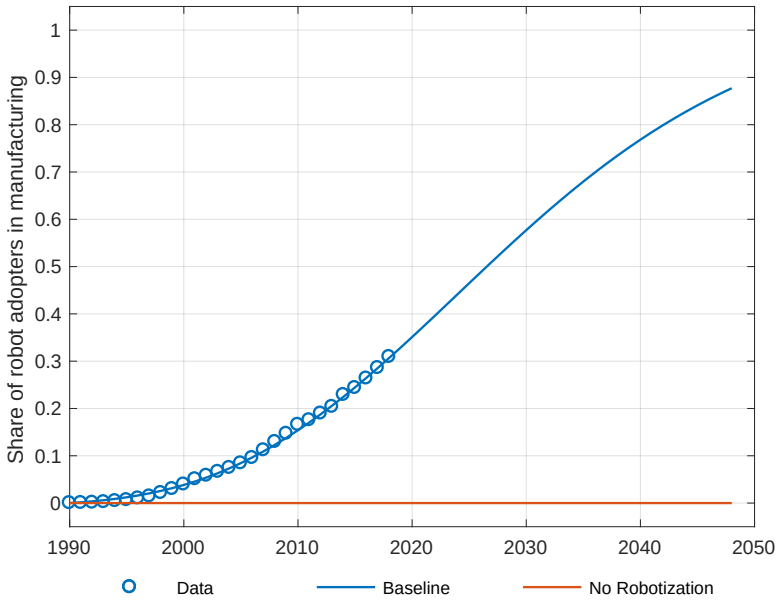
6.1.1 Solving the Dynamic General Equilibrium

The path of wages is the key endogenous variable that links the firm and worker decisions in general equilibrium. I solve for the general equilibrium wage schedule using a shooting algorithm adapted from Lee (2005). The procedure boils down to guessing a path of wages and manufacturing price indices, solving the dynamic programs related to the robot adoption decision of firms and the occupational choice problem of workers, simulating the economy forward using the firm and worker policy functions, and then using the firm's static labor demand functions to find the vector of wages that clear labor markets period-by-period. This algorithm iterates until convergence in the path of wages and the distributions of firm and worker states. Appendix F.3 details each step of the equilibrium solution algorithm.

6.2 The Distributional Impact of Industrial Robots

This section turns to the key question posed in this paper by asking how the distribution of worker earnings would have looked if industrial robots had not arrived. To evaluate this counterfactual, I solve the general equilibrium under a path of prohibitively high adoption costs ($c_t^R = \infty$). I then compare the results to the equilibrium under the baseline adoption cost schedule estimated in Section 4. The simulations assume that the arrival of industrial robot technology around 1990 came as a surprise to agents in the economy, but that firms and workers from that point on perfectly foresee the path of robot adoption costs. The robot diffusion curve in Figure 5 shows that if robot adoption had been infinitely costly ("No Robotization"), then robot technology would not have diffused at all.

Figure 5: Robot Diffusion Curve

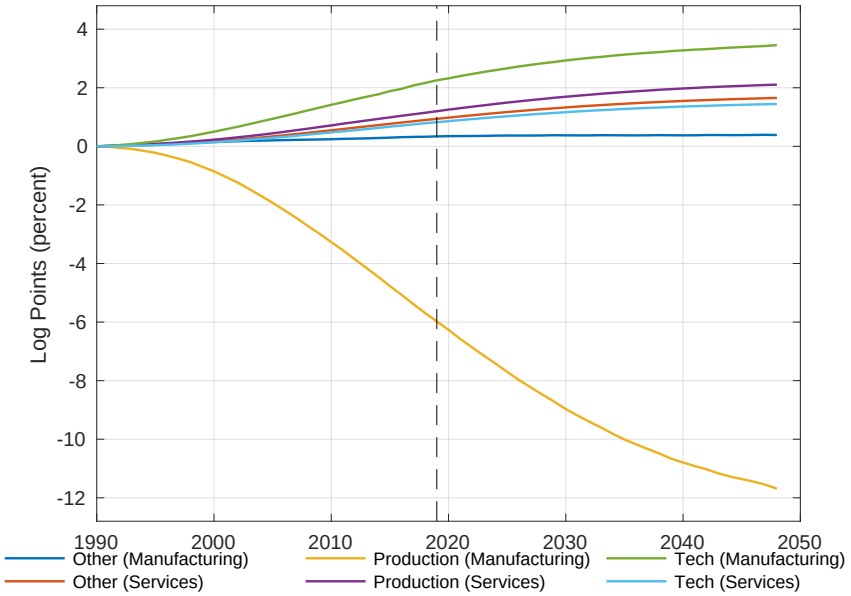


The equilibrium effects of industrial robots depend not only on the direct impact of

firm robot adoption estimated in Figures 1 and 2 but also on several indirect effects that are not identified in micro-level diff-in-diff regressions. The indirect effects include the extent to which the expansion of robot adopters crowds out non-adopter firms in product and labor markets as well as the ability of workers to reallocate across occupations in response to equilibrium wage pressures from robot diffusion. The general equilibrium model captures these indirect effects by combining the structurally estimated behavior of firms and workers with internal consistency constraints imposed by equilibrium conditions on product and labor markets.

Figure 6 shows the impact of industrial robots on real wages in different occupations. Industrial robots have increased average real wages by 0.8 percent in Denmark but with substantial distributional consequences. Production workers employed in manufacturing are the big losers from industrial robots, as their real wages are 6 percent lower today due to robots. Tech workers employed in manufacturing earn 2.3 percent higher real wages today due to industrial robots, while the remaining occupations have gained between 0.3 and 1.2 percent from robots. While the real wage loss for production workers in manufacturing is substantial, it is important to keep in mind that the occupation only constitutes around 3 percent of total employment in Denmark.

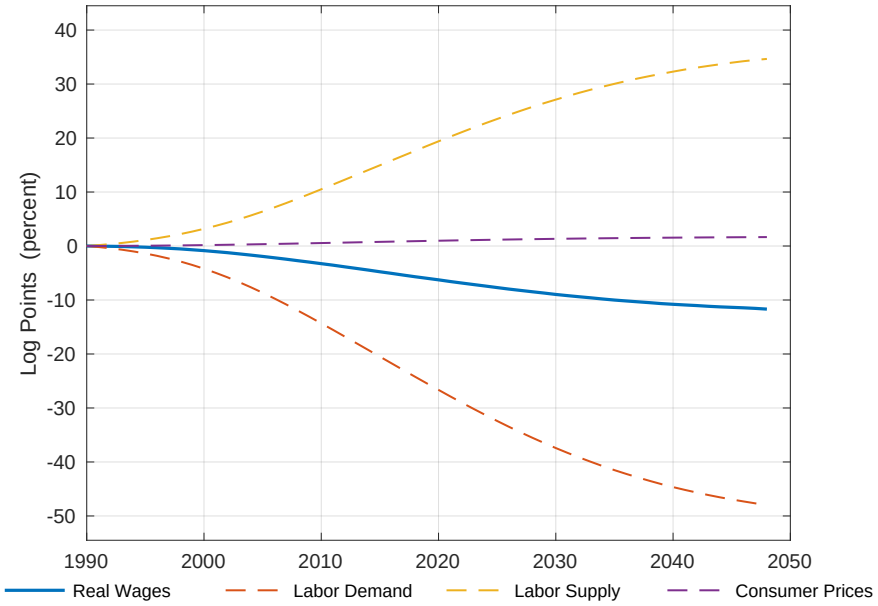
Figure 6: Real Wage Effects of Industrial Robots
(Weighted Average in 2019: +0.76 percent)



To understand the general equilibrium forces driving the real wage outcomes, Figure 7 decomposes the manufacturing production real wage effect into *labor demand* effects from robot adoption, *consumer price* effects from pass through of lower robot production costs, and *labor supply* effects from occupational reallocation of workers (changing the relative scarcity of labor across occupations).

As the decomposition shows, the real wage loss of manufacturing production workers would have been a half-order of magnitude larger than the estimated effect if workers could not reallocate across occupations in response to robots. Appendix Figure G.2 confirms this finding by evaluating the impact of industrial robots with exogenous labor supply, thus shutting off the occupational choice block estimated in Section 5.1. Real wages of production workers employed in manufacturing would in that world have been 30 percent lower today due to industrial robots.

Figure 7: Decomposition of the Wage Effect for Manufacturing Production Workers



Notes: Labor demand effects are measured relative to the "Other Workers" occupation in the services sector.

Still, the labor supply and consumer price effects combined are not enough to overturn the negative labor demand effects of robot adoption from depressing real wages of production workers employed in manufacturing. The displacement effects identified in Figure 2(a) are in general equilibrium reinforced by two additional labor demand forces. First, the expansion of robot adopters crowds out activity in non-adopter firms through the stealing of output markets. Second, the complementarity between occupations in manufacturing production (estimated in Section 4.1) means that firms spend a smaller fraction of their wage bill on production workers when they become less expensive.

Interestingly, among workers in the service sector, Figure 6 shows that production workers have experienced the largest real wage gain from robot adoption. This differential wage gain is a compensating differential for their excess risk of transitioning into production work in the manufacturing sector. In terms of expected lifetime earnings,

production workers are the group of service workers with the lowest gain from industrial robots.

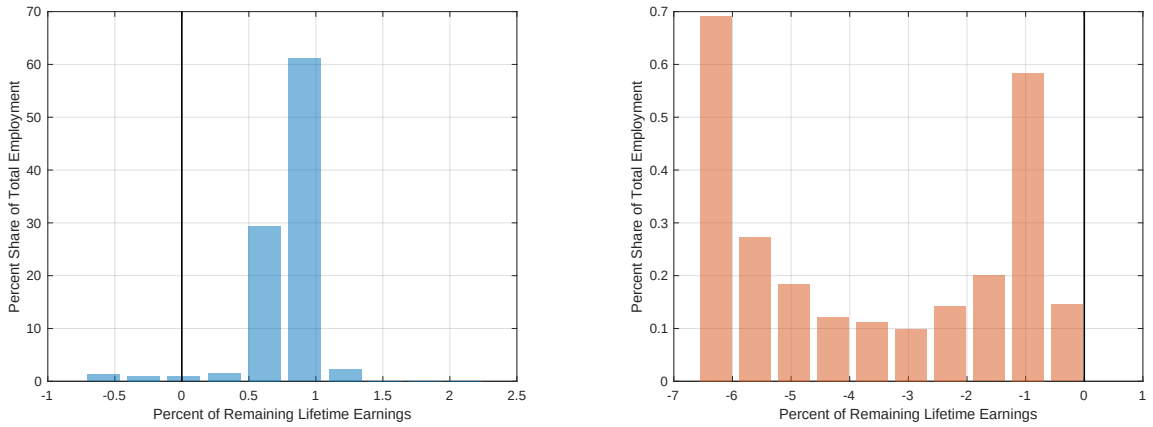
Finally, Figure 7 shows that more than half of the total consumer price gains from industrial robots have been realized already, even though only 30 percent of manufacturing firms have adopted robots. This finding reflects that the estimated model captures the fact that firms with larger efficiency gains from robot adoption (that is, firms that can better scale up production to take advantage of industrial robots) are the ones that adopt robots first.

Due to the possibility that workers can reallocate across occupations, the real wage effects in Figure 6 do not necessarily convert one-to-one into welfare effects for individual workers. The occupational reallocation margin opens an *option value* of being able to switch into occupations whose premiums rise as robots diffuse in the economy. As emphasized by Artuç et al. (2010), this option value source of worker welfare is not identified from static wage comparisons but is only captured once we factor in the dynamic occupational switching behavior observed over an individual's working life.

Figure 8 shows the impact of industrial robots on the welfare of workers in 2019. Panel (a) shows that 90 percent of workers have gained between 0.5 and 1 percent of lifetime earnings from the arrival of industrial robots. Yet, Panel (b) shows that the – considerably smaller – group of production workers employed in manufacturing have lost between 0 and 6 percent of lifetime earnings from robots.

Figure 8: Welfare Effects for Workers in 2019

(Average: +0.85 percent)



(a) All Workers Excl. Manufacturing Production

(b) Manufacturing Production Workers

Figure 9: Welfare Effects for Manufacturing Production Workers in 2019

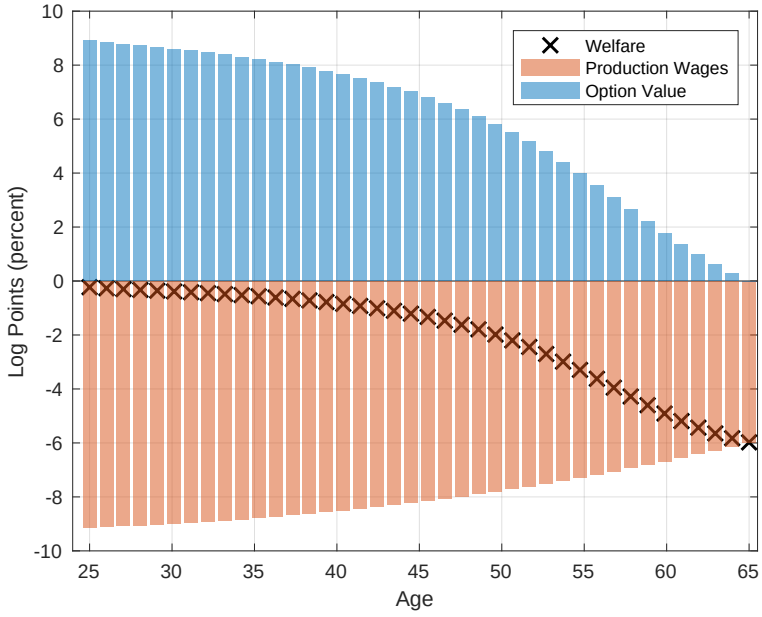
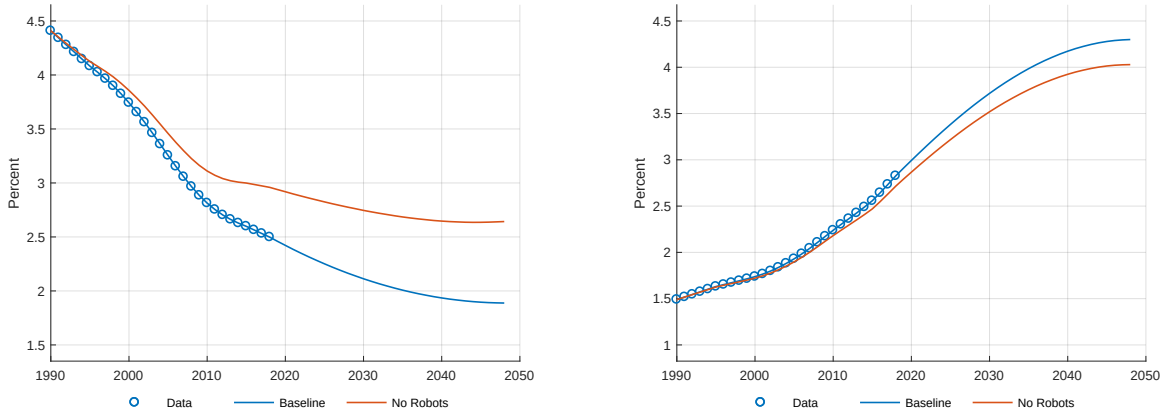


Figure 9 shows that the welfare losses in Figure 8(b) are concentrated on older workers. Younger production workers, with less specific skills and a long career ahead of them, are less affected by the arrival of industrial robots, as wage losses in their current occupation are offset by gains in the option value of switching into occupations whose premiums

rise as robots diffuse in the economy.

The flip side of the labor supply responses found in Figure 7 is that industrial robots have contributed to employment polarization as documented in Autor and Dorn (2013) and Goos et al. (2014). Figure 10 shows that industrial robots can account for 25 percent of the fall in the employment share of manufacturing production workers and 8 percent of the rise in the employment share of tech workers in manufacturing since 1990.

Figure 10: The Effect of Industrial Robots on Employment Shares



(a) Production Workers in Manufacturing

(b) Tech Workers in Manufacturing

To recapitulate, the estimates presented in this section are based on a general equilibrium model that has been validated on event studies of firm robot adoption, the observed diffusion of industrial robots, and worker transitions across occupations. The quantitative importance of the estimated general equilibrium responses to industrial robots warrants caution when comparing estimates from this section to findings in the reduced-form literature. For example, the conclusion from Figure 6 that industrial robots have increased average real wages may at first sight seem at odds with the finding in Acemoglu and Restrepo (2020) that robots depress wages in local labor markets. Before drawing such a comparison, however, it is important to keep in mind that Figure 6 takes into account the

general equilibrium consumer price and input-output linkage effects of industrial robots. For example, insofar as consumer price effects spill over across local labor markets, these contributions to real wages will be differenced out in the empirical strategy adopted in Acemoglu and Restrepo (2020). In fact, the average real wage gain estimated in Figure 6 flips to a loss if I omit the consumer price effect of industrial robots. Furthermore, to the extent that some of the positive contributions to the service sector through input-output linkages extend beyond commuting zones, these effects will also be differenced out in a diff-in-diff analysis of local labor markets. In fact, my estimate for the average wage effect in the manufacturing sector aligns well with the 0.4 percent loss estimated in Acemoglu and Restrepo (2020).

In summary, I take the estimates presented in this section as complementary to existing reduced-form studies of industrial robots by highlighting the quantitative importance of general equilibrium effects that are not easily identified by reduced-form empirical strategies. In particular, I show the quantitative relevance of an occupational switching feedback mechanism that has been emphasized in the literature on international trade and labor market dynamics (Dix-Carneiro, 2014; McLaren, 2017; Traiberman, 2019). Although the labor supply responses are not strong enough to overturn the negative labor demand effects from depressing the real wages of manufacturing production workers, I find that the wage losses would have been a half-order of magnitude larger if workers could not reallocate across occupations. A speculative hypothesis is that the generous retraining subsidies offered in the Danish system of active labor market policies could be an underlying driver of the quantitative importance of the estimated occupational reallocation feedback response.

6.3 Policy Counterfactuals: The Dynamic Incidence of a Robot Tax

As a final counterfactual experiment, I now turn to evaluating the impact of a robot tax. The European Parliament voted in 2017 on a proposal to tax the use of robotics. The robot tax was motivated as a way to slow down the speed of robot adoption to give the economy more time to adjust to the new technology.[19]

I tax the schedule of robot adoption costs c_t^R to inform this policy counterfactual. To be clear, the undistorted equilibrium of the model is efficient (except for markups in product markets), but the robot tax could be motivated by distributional concerns.[20] In particular, Section 6.2 identified a group of production workers employed in manufacturing who have clearly lost from the use of industrial robots. A key policy question is how costly (in terms of lost economic efficiency) it is to insulate these production workers by taxing the further adoption of industrial robots. The answer to this question depends on several behavioral elasticities estimated from the micro data, including the sensitivity of firm robot adoption with respect to adoption costs (Section 4.4.2) as well as the ability of workers to switch occupations in response to robots (Section 5.1). I use the estimated general equilibrium model to quantify the distributional implications of a robot tax and to evaluate its impact on aggregate economic activity.

To map out the potential policies, I evaluate both a temporary and a permanent tax, each of 30 percent. The policies are announced and implemented in 2019, and the tempo-

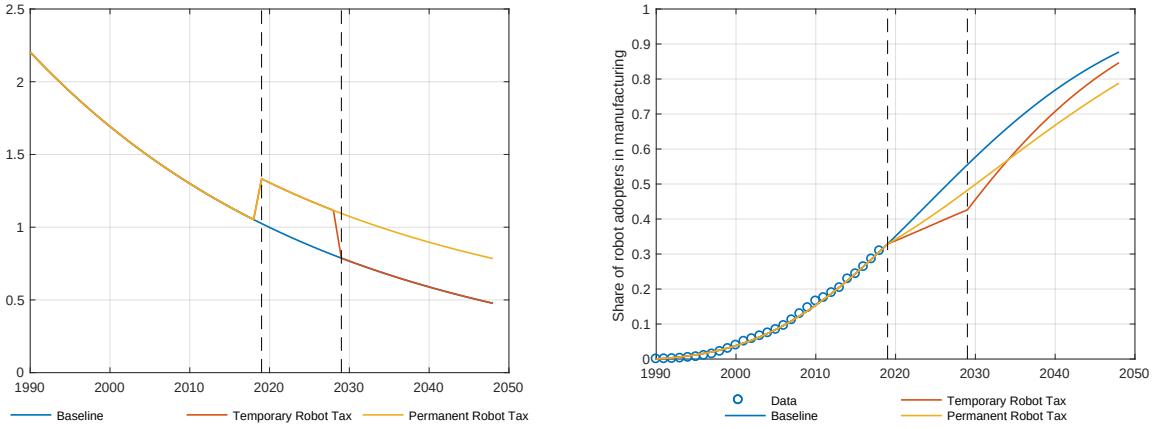
[19]The proposal was ultimately voted down by the European Parliament but the idea of taxing robots to mitigate labor market polarization remains popular among public figures from Bill Gates (Quartz, 2017) to congresswoman Alexandria Ocasio-Cortez (Market Watch, 2019).

[20]The *production efficiency* result of Diamond and Mirrlees (1971) establishes that it is always optimal to maintain production efficiency insofar as linear commodity taxes are available. Costinot and Werning (2018) derive sufficient-statistic formulas for optimal technology taxes when a non-linear income tax schedule is the only alternative policy instrument.

rary tax is put in place for 10 years. Figure 11(a) shows the path of robot adoption costs under the tax policies. I assume that a robot tax in Denmark does not alter the pre-tax price for robots which is determined on world markets.

Panel (b) of Figure 11 shows the first key result from the robot tax counterfactuals: The temporary tax is more effective in slowing down the diffusion of industrial robots while it is put in place. With the temporary tax, only 43 percent of manufacturers will have adopted robots by 2029, compared to 48 percent with the permanent tax and 56 percent in the baseline scenario. The larger short-term effects of the temporary tax reflect the forward-looking nature of adoption, where firms foresee that the robot tax will expire and postpone adoption until then. The flip side of these delays is that the adoption of robots accelerates beyond its baseline speed after the temporary tax expires in 2030.

Figure 11: Robot Tax Counterfactuals

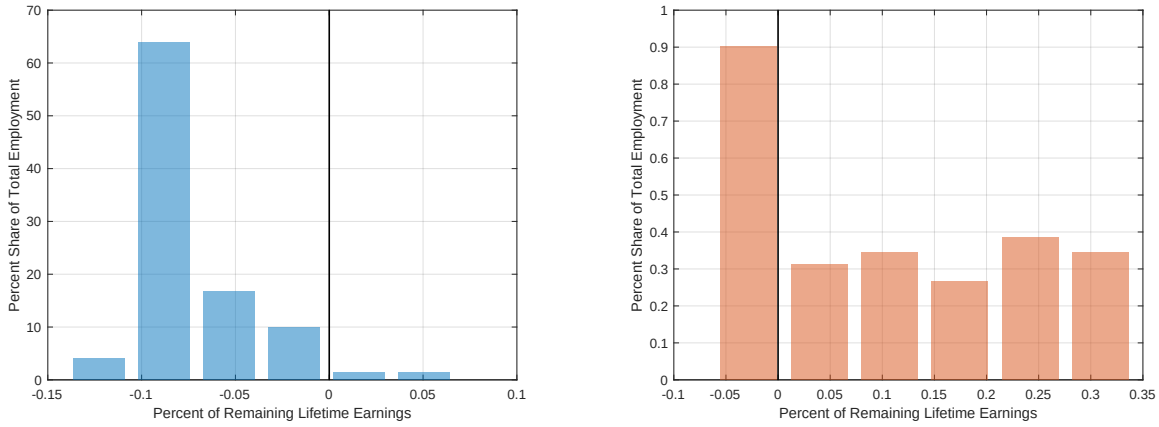


(a) Robot Adoption Costs

(b) Robot Diffusion Curve

Figure 12 shows how the temporary robot tax affects the welfare of workers in 2019. The temporary tax lowers average welfare by 0.05 percent of lifetime earnings but benefits a group of older production workers employed in manufacturing by 0.2 to 0.3 percent.

Figure 12: The Impact of a Temporary Robot Tax on the Welfare of Workers in 2019
(Average: -0.054 percent)



(a) All Workers Excl. Manufacturing Production (b) Manufacturing Production Workers

Table 5 shows how the burden of the robot taxes falls on workers and firms in the economy. Measured in presented discounted terms, the robot taxes redistributes a total of 0.01 to 0.02 percent of GDP to production workers currently employed in manufacturing at the expense of a total welfare loss for workers of around 1 percent of GDP. These welfare losses reflect foregone efficiency gains from underinvestment in robot technology. Put differently, for the robot taxes to enhance social welfare, one needs to value production workers in manufacturing 50 to 100 times higher than the average worker.

The temporary robot tax creates welfare losses per dollar of tax revenue collected that are considerably larger than those of the permanent robot tax. These larger relative efficiency losses of the temporary tax is a direct consequence of the investment delays observed in Panel (b) of Figure 11: The intertemporal shifting of robot adoption out of the temporary policy window creates misallocation without raising tax revenues. In particular, if firm adoption behavior did not respond to the robot tax ("Mechanical Effect" in

Table 5), the temporary robot tax would generate 133 percent more revenues, while revenues from the permanent tax would be only 17 percent higher.

The robot taxes do, however, generate substantial amounts of tax revenue, whose burdens are primarily borne by manufacturing firms. As Table 5 shows, the tax revenues are sufficient to make all workers better off from the robot taxes, insofar as the revenues can be rebated appropriately and the planner does not care about firm profits. One should be cautious about drawing such a conclusions, however, as I do not model firms' entry decisions. If the robot taxes would cause some manufacturing firms to go out of business, these profit losses would be passed on to lower worker welfare.

Table 5: Robot Tax Incidence
(Discounted Present Values in Percent of GDP in 2019)

	Temporary Tax	Permanent Tax
Workers	-1.21	-1.00
Workers in 2019	-0.62	-0.47
– Manufacturing Production	0.02	0.01
Future Workers	-0.59	-0.53
Tax Revenues	2.39	9.41
Mechanical Effect	5.57	11.02
Behavioral Effect	-3.18	-1.61
Profits (excl. predatory externalities)	-4.14	-10.58

Notes: Workers represent compensating variations; see Appendix G.1.1 for details. *Profits (excl. predatory externalities)* represent the effect on manufacturing firm values (Equations (7)-(8)) in 2019, holding constant pecuniary externalities of robot adoption in output markets; see Appendix G.2.1 for details. *Mechanical Effect* is the tax revenues collected if robot adoption did not respond to the tax.

In calculating the effects on firm profits in Table 5, I exclude so-called predatory investment externalities. Predatory investments refer to the pecuniary externality where robot adopters do not internalize that parts of the profit gain from robots come from stealing markets shares of competitor firms.[21] By internalizing this predation effect, a robot tax has

[21]The implications of predatory investments have been studied extensively in the theoretical industrial

the possibility to increase aggregate profits of firms. To focus on the key equity-efficiency trade-off for workers, I hold the predatory externalities out of the baseline incidence calculations, and instead relegate their analysis to Appendix G.2.1.

To sum up, even though the temporary tax achieves the goal of delaying the diffusion of industrial robots, this analysis shows that the policy is an ineffective and relatively costly way to redistribute income to production workers employed in manufacturing.

7 Conclusion

This paper makes two methodological contributions in order to study the distributional impact of industrial robots. First, I develop a dynamic firm model that can rationalize the selection into and reduced-form responses to robot adoption. Second, I model both firm and worker dynamics in general equilibrium. I use administrative data that link workers, firms, and robots in Denmark to structurally estimate a dynamic general equilibrium model that can account for event studies of firm robot adoption, the observed diffusion of industrial robots, and worker transitions in the labor market. The model fits the labor demand responses to robot adoption but also takes into account how production efficiency gains from robots are passed through to lower consumer prices as well as the ability of workers to reallocate between occupations in response to industrial robots.

Having validated the model using overidentification checks, I use it to estimate the distributional impacts of industrial robots. I find that industrial robots have increased average real wages by 0.8 percent but with substantial distributional consequences. At

organization literature, including Dixit (1980) and Spence (1986).

the ends of the spectrum, I find that production workers employed in manufacturing have lost 6 percent in real wages while tech workers have gained 2.3 percent.

The model captures worker heterogeneity in exposure to robot diffusion across occupation, industry, tenure, skill, and age of workers but abstracts from the possibility that robot adoption could differentially affect incumbent workers in the adopting firm. Using matched worker-firm-robot datasets to collect evidence on how firm robot adoption affects incumbent workers, as in Bessen et al. (2019), represents a promising avenue of future empirical research. Introducing such firm-specific wage or displacement effects into the general equilibrium framework developed in this paper without breaking the block separability that keeps the model computationally tractable is an important avenue of future theoretical research. I lay out one such model extension in the appendix of this paper.

I believe that the quantitative framework developed in this paper can be applied to studying the labor market impacts of other pressing technologies. For example, what will be the consequences when 1.3 million US truck drivers are expected to compete with self-driving vehicle technology by 2026 (Council of Economic Advisers, 2016)? The quantitative experiments conducted in this paper highlight that the ability of workers to switch occupations is crucial for how a new technology can affect the distribution of earnings. In Humlum (2019), I find that retraining subsidies can be an effective tool to help workers transition across occupations in the labor market. These findings may help policymakers navigate in an era of rapid technological change.

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 (2018a). Demographics and Automation. Technical report, National Bureau of Economic Research.

Acemoglu, D. and P. Restrepo (2018b). Modeling Automation. *AEA Papers and Proceedings 108*, 48–53.

Acemoglu, D. and P. Restrepo (2019). Automation and New Tasks: How Technology Displaces and Reinstates Labor. *Journal of Economic Perspectives 33*(2), 3–30.

Acemoglu, D. and P. Restrepo (2020). Robots and Jobs: Evidence From US Labor Markets. *Journal of Political Economy 128*(6), 2188–2244.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification Properties of Recent Production Function Estimators. *Econometrica 83*(6), 2411–2451.

Adda, J. and R. Cooper (2003). *Dynamic Economics: Quantitative Methods and Applications*. MIT press.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arcidiacono, P. and R. A. Miller (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity. *Econometrica 79*(6), 1823–1867.

Artuç, E., S. Chaudhuri, and J. McLaren (2010). Trade Shocks and Labor Adjustment: A Structural Empirical Approach. *American Economic Review 100*(3), 1008–45.

Ashournia, D. (2017). Labour Market Effects of International Trade When Mobility Is Costly. *The Economic Journal 128*(616), 3008–3038.

Asker, J., A. Collard-Wexler, and J. De Loecker (2014). Dynamic Inputs and Resource (Mis)Allocation. *Journal of Political Economy 122*(5), 1013–1063.

Autor, D. and D. Dorn (2013). The Growth of Low-Skill Service Jobs and the Polarization of the Us Labor Market. *American Economic Review 103*(5), 1553–97.

Backus, D. K., P. J. Kehoe, and F. E. Kydland (1992). International Real Business Cycles. *Journal of Political Economy 100*(4), 745–775.

Bernard, A. B., T. C. Fort, V. Smeets, and F. Warzynski (2018). Heterogeneous Globalization: Offshoring and Reorganization. Technical report, Working Paper, Dartmouth College.

Bernard, A. B., V. Smeets, and F. Warzynski (2017). Rethinking Deindustrialization. *Economic Policy 32*(89), 5–38.

Berry, S., A. Gandhi, and P. Haile (2013). Connected Substitutes and Invertibility of Demand. *Econometrica 81*(5), 2087–2111.

Bessen, J. E., M. Goos, A. Salomons, and W. Van den Berge (2019). Automatic Reaction – What Happens to Workers at Firms That Automate? *Boston Univ. School of Law, Law and Economics Research Paper*.

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica 77*(3), 623–685.

Caliendo, L., M. Dvorkin, and F. Parro (2019). Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica 87*(3), 741–835.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics 36*(S1), S13–S70.

Costinot, A. and I. Werning (2018). Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation. Technical report, National Bureau of Economic Research.

Council of Economic Advisers (2016). Artificial Intelligence, Automation, and the Economy. *Executive Office of the President*, 18–19.

Dauth, W., S. Findeisen, J. Suedekum, N. Woessner, et al. (2018). Adjusting to Robots: Worker-Level Evidence. Technical report.

Delvaux, M. (2016). Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics. *European Parliament Committee on Legal Affairs*.

Diamond, P. A. and J. A. Mirrlees (1971). Optimal Taxation and Public Production I: Production Efficiency. *The American Economic Review 61*(1), 8–27.

Dix-Carneiro, R. (2014). Trade Liberalization and Labor Market Dynamics. *Econometrica 82*(3), 825–885.

Dixit, A. (1980). The Role of Investment in Entry-Deterrence. *The Economic Journal 90*(357), 95–106.

Doraszelski, U. and J. Jaumandreu (2013). R&D and Productivity: Estimating Endogenous Productivity. *Review of Economic Studies 80*(4), 1338–1383.

Doraszelski, U. and J. Jaumandreu (2018). Measuring the Bias of Technological Change. *Journal of Political Economy 126*(3), 1027–1084.

Dubin, J. A. and D. L. McFadden (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, 345–362.

Fort, T. C., J. R. Pierce, and P. K. Schott (2018). New Perspectives on the Decline of Us Manufacturing Employment. *Journal of Economic Perspectives 32*(2), 47–72.

Goos, M., A. Manning, and A. Salomons (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review 104*(8), 2509–26.

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

Griliches, Z. (1957). Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica*, 501–522.

Hawkins, W., R. Michaels, and J. Oh (2015). The Joint Dynamics of Capital and Employment at the Plant Level. *Working Paper, Yale University*.

Hotz, V. J. and R. A. Miller (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models. *The Review of Economic Studies 60*(3), 497–529.

Humlum, A. (2019). Retraining and Occupational Choice: Evidence from a Natural Experiment. *Working Paper, Princeton University*.

Hummels, D., R. Jørgensen, J. Munch, and C. Xiang (2014). The Wage Effects of offshoring: Evidence from Danish Matched Worker-Firm Data. *American Economic Review 104*(6), 1597–1629.

Imbens, G. and J. Wooldridge (2007). Estimation of Average Treatment Effects Under Unconfoundedness. *What's New in Econometrics, Lecture Notes 1*.

International Federation of Robotics (2018). World Robotics Report 2018.

King, R. G. and S. T. Rebelo (1999). Resuscitating Real Business Cycles. *Handbook of Macroeconomics 1*, 927–1007.

Koch, M., I. Manuylov, and M. Smolka (2019). Robots and Firms. *CESifo Working Paper*.

Lee, D. (2005). An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice. *International Economic Review 46*(1), 1–34.

Lee, D. and K. I. Wolpin (2006). Intersectoral Labor Mobility and the Growth of the Service Sector. *Econometrica 74*(1), 1–46.

Levinsohn, J. and A. Petrin (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies 70*(2), 317–341.

Market Watch (2019). Bill Gates Finds an Ally in Washington for His Idea to Tax Robots: Alexandria Ocasio-Cortez. *Market Watch. March 11, 2019*.

Marschak, J. and W. H. Andrews (1944). Random Simultaneous Equations and the Theory of Production. *Econometrica*, 143–205.

McLaren, J. (2017). Globalization and Labor Market Dynamics. *Annual Review of Economics 9*, 177–200.

Olley, G. S. and A. Pakes (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 1263–1297.

Quartz (2017). The Robot That Takes Your Job Should Pay Taxes, Says Bill Gates. *Quartz Magazine. February 17, 2017*.

Raval, D. R. (2019). The Micro Elasticity of Substitution and Non-Neutral Technology. *The RAND Journal of Economics 50*(1), 147–167.

Region Syddanmark (2017). Robotter og Automatisering – Styrkepositioner, Udfordringer og Udviklingspotentiale. Technical report, Region Syddanmark.

Rubin, D. B. (1990). Formal Mode of Statistical Inference for Causal Effects. *Journal of Statistical Planning and Inference 25*(3), 279–292.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 999–1033.

Spence, M. (1986). Cost Reduction, Competition and Industry Performance. In *New Developments In the Analysis of Market Structure*, pp. 475–518. Springer.

Stokey, N. L. and R. Lucas (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

Tauchen, G. (1986). Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions. *Economics Letters 20*(2), 177–181.

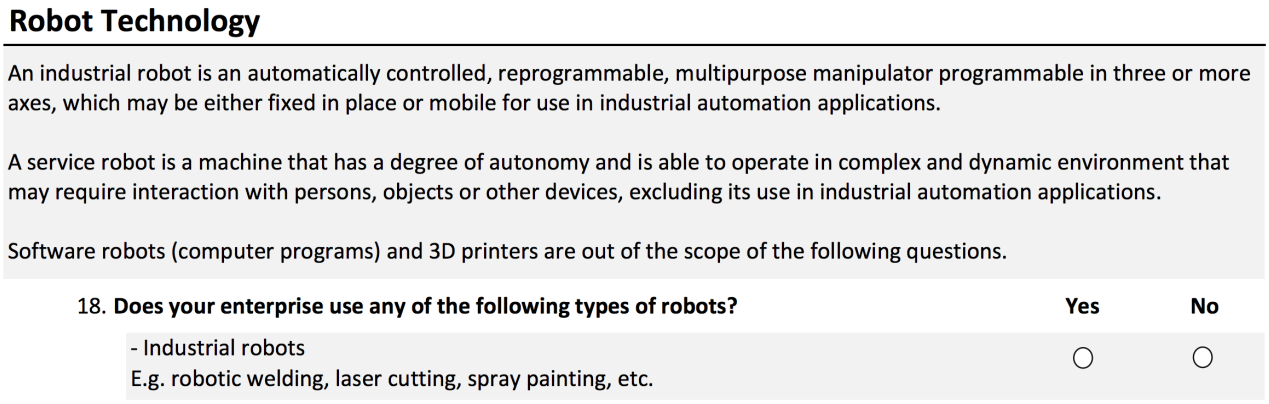
Traiberman, S. (2019). Occupations and Import Competition: Evidence From Denmark. *American Economic Review 109*(12), 4260–4301.

A Data

A.1 Robot Adoption Firm Survey

Statistics Denmark conducts annually a technology adoption survey of firms in Denmark (IT usage in Danish enterprises, VITA). The survey is prepared in collaboration with the Danish Business Authority as a supplement to Eurostat's technology survey. In 2018, the survey included a question on the use of industrial robots. The survey sampled 3,954 firms from the population of 16,465 private non-agricultural, non-financial firms with more than 10 employees. The response rate was 97 percent. Figure A.1 shows the questionnaire on industrial robot usage. Out of the survey respondents, a total of 473 firms answered 'yes' to using industrial robots in production.

Figure A.1: Firm Questionnaire on Firm Robot Adoption



Robot Technology

An industrial robot is an automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes, which may be either fixed in place or mobile for use in industrial automation applications.

A service robot is a machine that has a degree of autonomy and is able to operate in complex and dynamic environment that may require interaction with persons, objects or other devices, excluding its use in industrial automation applications.

Software robots (computer programs) and 3D printers are out of the scope of the following questions.

18. **Does your enterprise use any of the following types of robots?**	**Yes**	**No**
- Industrial robots E.g. robotic welding, laser cutting, spray painting, etc.	○	○

A.2 Firm Customs Records

The firm customs records are organized in the Foreign Trade Statistics Register (UHDI) at Statistics Denmark. For each firm in each year 1993-2015, I have imports disaggregated

by origin and 6-digit Harmonized System product code. One of these codes identifies "847950 Industrial Robots".[22] Industrial robots are heavily imported goods in Denmark (import share of 95 percent according to calculations in Section A.4), making customs records a valuable source of information on the adoption of industrial robots. The main challenge in using the customs records is that a substantial share of machinery is imported through domestic distributors. Table A.3 develops a procedure for identifying robot imports done by final adopters.[23] Starting from the population of robot imports, I

1. *Pre-data coverage*: Restrict the sample to firms who are active three years before the import event. This condition is necessary for conducting the adoption event studies.
2. *Exclude wholesalers*: Exclude the one-digit industry code "5 Commerce".
3. *Exclude integrators*: Exclude 6-digit industry codes contained in a comprehensive list of robot integrators in 2018.[24]
4. *Survey-validated adoptions*: Restrict the sample to 6-digit industries with a validation share in the robot adoption survey (VITA) of minimum 50 percent. The validation share is defined as the fraction of robot importers that in the robot adoption survey report that they use industrial robots.[25]

[22]Fort et al. (2018) use this product code to collect descriptive evidence on robot importers in the United States.

[23]I thank several industry experts for helpful inputs into developing this sample selection procedure, including Søren Peter Johansen (Technology Manager at the Danish Technological Institute, Robot Technology), Bo Hanfgarn Eriksen (Region Syddanmark), Per Rasmussen (BILA Robotics), and Martin Jespersen (Odense Robotics).

[24]List of industry codes excluded: 51.60.00 Wholesale of machinery and equipment, 30.00.09 Manufacture of computer equipment, electric motors, etc., 29.40.09 Manufacturing of industrial machinery, 29.00.00 Manufacture of multi-purpose machines, 28.10.09 Manufacture of metal building materials. The list of robot integrators was developed by RoboCluster and Odense Robotics for the report Region Syddanmark (2017). I thank Bo Hanfgarn Eriksen at Region Syddanmark for providing the list.

[25]List of industry codes included: 33.00.00 Manufacture of medical equipment, 30.00.09 Manufacture of computer equipment, electric motors, 29.30.00 Manufacture of agricultural machinery, 29.20.00 Manufacture of general purpose machinery, 29.10.00 Manufacture of ship engines, compressors, etc., 28.60.09 Manufacture of hand tools, metal packaging, etc., 28.10.09 Manufacture of metal building materials, 26.30.09

5. *Single production establishment*: Restrict the sample to firms that only have a single establishment employing more than three workers in the year prior to robot adoption. This condition avoids dilution of the robot adoption effect in multi-plant firms (robot adoption happens at the plant level, but customs forms are filled out at the firm level).

The sample selection criteria exclude many of the robot import observations. For the sake of sustaining power in the statistical analysis, I use the 4-digit product code that includes industrial robots (HS 8479), as also done in Acemoglu and Restrepo (2018a).

Table A.1: Identifying Robot Adoption in Customs Records

Step	Sample at End of Step		
	Imports (million USD)	Import events (firm-year)	Firms
Raw imports	3291.9	14355	4839
1. Pre-data coverage	1457.7	5936	2594
2. Exclude wholesalers	826.5	2016	1048
3. Exclude integrators	535.0	1375	754
4. Survey-validated industries	247.6	776	416
5. Single production establishment	91.1	454	293

A.3 Comparison of Data Sources on Robot Adoption

This section compares three data sources on robot adoption against each other: the robot adoption firm survey (Section A.1), the firm customs records (Section A.2), and the International Federation of Robotics (IFR) statistics. While the IFR statistics have been the main source of data for the existing papers on robotization (Acemoglu and Restrepo, 2020;

Brick, cement, and concrete industries, 25.00.00 Rubber and plastic industry, 24.40.00 Pharmaceutical Industry, 24.30.09 Manufacture of paints, soaps, cosmetics, etc., 24.10.09 Manufacture of chemical raw materials, 20.00.00 Wood industry, 15.89.09 Other food industry.

Dauth et al., 2018; Graetz and Michaels, 2018), the present paper is the first to use the adoption survey and the Danish customs records to study robot adoption. Section A.3.1 compares the industry representation across the three data sources, and Section A.3.2 examines how the time series of robot adoption compare in the different datasets.

A.3.1 Cross-Sectional Comparisons

Table A.2 shows that the industry composition of robot adoption in the micro data used in the present paper align well with the statistics compiled by the International Federation of Robotics. The data sources agree that industrial robots are a manufacturing technology, and that the metal, chemical, and plastic industries have been the main drivers of robot adoption.

Table A.2: Robot Adoption Across Industries: Comparison of Data Sources

	Data Sources		
	Robot Survey (StatDK)	Robot Stock (IFR)	Robot Imports (Customs)
Share in Total Adoptions (%)			
Manufacturing	79.1	85.9	83.5
Share in Manufacturing Adoptions (%)			
Food and beverages	7.2	18.3	7.2
Textiles	1.1	2.8	0.0
Wood and furniture	6.4	4.7	3.5
Paper	2.2	1.4	0.0
Plastic and Chemicals	14.0	22.0	32.3
Glass, stone, minerals	5.0	3.7	1.9
Metal	51.1	34.1	31.7
Electrical and Electronics	10.9	8.4	23.5
Automotives and vehicles	1.9	4.7	0.0

Notes: "Robot Survey" indicates the share in total firm robot adopters. "Robot Stock" specifies the share in total robot stock. "Robot Imports" is the share in total firm robot import events (firm-year observations). Robot Imports represents the 454 adoption events identified in Table A.1.

A.3.2 Time Series Comparisons

The IFR statistics and the customs records each contain a time series dimension allowing me to compare how robot adoption has evolved according to the two data sources. Figure A.2 shows that total robot imports in Denmark (from custom records) have closely tracked the total number of robot installments (IFR statistics) since the 1990s.[26]

Figure A.2: Robot Adoption Across Time: Comparison of Data Sources



A.4 Calculation of Robot Import Absorption Share

This section calculates the import share in robot adoption using micro data on re-exports and domestic producers. The robot import share is defined as

$$\text{Import Absorption Share}_t = \frac{\text{Imports Absorbed}_t}{\text{Imports Absorbed}_t + \text{Production for Domestic Absorption}_t}$$

[26]The time series are normalized to 1 in 2010. This normalization implies a robot unit price of $58,000, which falls within the range of common list prices for industrial robots ($50,000 to $100,000 according to the International Federation of Robotics (2018)).

I measure import absorption by summing over firms' robot imports, netting out their robot exports and transit trade. To measure the domestic supply (production for domestic absorption), I leverage the high export-orientation of robot producers to impute the domestic sales of robots. In particular, I use the customs records for the exports of robot producers (list provided by industry experts) to calculate the share of robots in total sales of the firm. I then multiply this robot share with the firm's domestic sales to impute the firm's domestic sales of robot.[27] Table A.3 shows that the robot import share has averaged 94.9 percent from 1993 to 2015.

Table A.3: Import Share in Robot Investments, Denmark 1993-2015 (percent)

Average	1993-2004	2005-2015
94.9	98.5	90.9

A.5 Measuring Domestically Sourced Robot Adoptions

This section describes how I supplement the customs records to measure robot adoption done through domestic distributors. I first use the representative robot adoption firm survey (VITA) conducted by Statistics Denmark; see Appendix A.1 for details. The survey provides a snapshot of which firms use industrial robots in 2018, regardless of whether the firms have imported their robots directly or have relied on a domestic distributor. From the adoption survey, I can directly calculate that 31 percent of manufacturing firms have adopted robots (last data point in Figure 3(a)) and that these adopters represent 54 percent of manufacturing sales (Figure 4).

For the time series of robot adoption, I use the International Federation of Robotics

[27] Measuring the domestic supply (production for domestic absorption) is complicated by the fact that product code breakdowns of domestic sales in general do not exist.

(IFR) statistics on the stock of industrial robots in Danish manufacturing over time (the data source of Acemoglu and Restrepo (2020) and Graetz and Michaels (2018)). Assuming that the robot stock per adopter firm is constant over time, I can use the IFR time series to extend the number of robot adopters observed in 2018 back in time (Figure 3(a)). As a robustness check, I verify that the robot imports data imply the same evolution in total robot adoption over time.

A.6 Occupational Classification

I build on the occupational classification developed by Bernard et al. (2017) to study worker tasks. The classification groups detailed four-digit ISCO codes into six categories: managers, tech workers, sales workers, support workers, transportation/warehousing, and line workers (mostly production). The classification is used in Bernard et al. (2018).

Table 2 shows that the robot adopters and match firms are balanced on these occupational categories prior to adoption. In the main analysis, I focus on the three occupations that are most relevant to industrial robots: tech workers, production workers, and other workers. Tech workers is the second category of the Bernard et al. (2017) classification, and includes skilled technicians, engineers, and researchers. Production workers is the intersection of the sixth category of Bernard et al. (2017) (line workers, mostly production) and the 1-digit ISCO88 code "7 Craft and Related Trades Workers." Production workers consist of manual production tasks from welding to assembly.

A.7 Stylized Facts on Firm Robot Adoption

A.7.1 Matching Procedure

This section describes the matching algorithm used in column 3 of Table 2. The procedure is structured as follows.

1. Pick a vector X_e to match exactly on, and a vector $X_d \in \mathbb{R}^K$ to distance match on.

2. For each adopter firm f, find non-adopter match firm g that

   (a) matches f exactly on X_e

   (b) has minimal Mahalanobis distance to f in X_d

$$\text{Match}_f = \arg \min_{g \in \{X_e(f) \cap \text{na}\}} (X_{dg} - X_{df})' \Sigma (X_{dg} - X_{df}),$$

   where Σ is the sample covariance matrix of X_d.

In my application, I match exactly (X_e) on industry (two-digit) and year $t-1$. Within each industry-year bin, I then distance match (X_d) on firm sales and production line wage bill shares (levels at $t-1$ and changes from $t-3$).

B A Model of Firm Robot Adoption

B.1 Task-Based Micro Foundation for the Production Function

This section provides a task-based micro foundation for the production function used in Section 3.[28] Consider a firm j operating the task-based production technology

$$Y_{jt} = \left(\int_0^{\mathcal{I}_{jt}} Y_{jt}(i)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}, \quad Y_{jt}(i) = z_{ojt}(i) X_{ojt}(i) \mathbf{1}_{\{i \in \mathcal{A}_{ojt}\}}, \tag{39}$$

where production tasks are indexed by i and factors (production workers, tech workers, intermediate inputs, etc.) are indexed by o. Let $\mathcal{A}_{jt} = \{\mathcal{A}_{1jt}, .., \mathcal{A}_{\mathcal{O}jt}\}$ denote an assignment of tasks to factors (a partition of the interval $[0, \mathcal{I}_{jt}]$). Conditional on such a task assignment, the firm has to allocate the time of each factor across its assigned tasks. The first-order conditions to this time allocation problem are

$$X_{ojt}(i) = \frac{z_{ojt}(i)^{\sigma-1}}{\int_{i \in \mathcal{A}_{ojt}} z_{ojt}(i)^{\sigma-1} di} X_{ojt} \quad \text{for } i \in \mathcal{A}_{ojt}, \tag{40}$$

where $X_{ojt} = \int_{i \in \mathcal{A}_{ojt}} X_{ojt}(i) di$ is the total units of factor o employed at firm j in year t. By inserting Equation (40) into Equation (39), I can now represent the firm's technology with

[28] The setup is inspired by Hawkins et al. (2015), who study the cost minimization problem of a plant operating a Ricardian task-based production technology where the assignment of productive factors to tasks is subject to a Calvo shock.

the production function

$$Y_{jt} = Y(X_{jt} \mid \mathcal{I}_{jt}, z_{jt}) = \left(\sum_{o \in \mathcal{O}} (z_{ojt} L_{ojt})^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}, \quad \text{with} \tag{41}$$

$$z_{ojt} = \frac{\int_{i \in \mathcal{A}_{ojt}} z_{ojt}(i)^{\sigma}_{ojt}}{\int_{i \in \mathcal{A}_{ojt}} z_{ojt}(i)^{\sigma-1}_{ojt}} \tag{42}$$

Following Acemoglu and Restrepo (2019), suppose that firm robot adoption may

1. Affect each factor's productivity in a given task (the *productivity* effect), $z_{jt}(i, R_{jt})$
2. Require tasks to be reassigned between factors (the *substitution* effect), $\mathcal{A}_{jt}(R_{jt})$
3. Create new tasks to be carried out in production (the *reinstatement* effect), $\mathcal{I}_{jt}(R_{jt})$.

I can then reformulate Equation (41) into a robot-contingent production function

$$Y_{jt} = Y(X_{jt} \mid R_{jt}, \omega_{jt}) = \left\{ \sum_{o \in \mathcal{O}} (z_{ojt} L_{ojt})^{\frac{\sigma-1}{\sigma}} di \right\}^{\frac{\sigma}{\sigma-1}}, \text{with} \tag{43}$$

$$z_{ojt} = \exp(\varphi_{ojt} + \gamma_{ojt} R_{ojt}) \tag{44}$$

$$\varphi_{ojt} = \log \frac{\int_{i \in \mathcal{A}_{ojt}(0)} z_{ojt}(i,0)^{\sigma}}{\int_{i \in \mathcal{A}_{ojt}(0)} z_{ojt}(i,0)^{\sigma-1}} \tag{45}$$

$$\gamma_{ojt} = \log \frac{\int_{i \in \mathcal{A}_{ojt}(1)} z_{ojt}(i,1)^{\sigma}}{\int_{i \in \mathcal{A}_{ojt}(1)} z_{ojt}(i,1)^{\sigma-1}} - \log \frac{\int_{i \in \mathcal{A}_{ojt}(1)} z_{ojt}(i,0)^{\sigma}}{\int_{i \in \mathcal{A}_{ojt}(0)} z_{ojt}(i,0)^{\sigma-1}} \tag{46}$$

Equations (43)-(46) provide a direct micro foundation of the production function used in Equation (1). The only parametric restrictions imposed in Equation (1) are that of homogeneous robot productivity effects, $\gamma_{jt} = \gamma$.[29]

[29]The micro foundation in Equations (43)-(46) provides insights into the task-based sources of treatment effect heterogeneity in robot adoption. A promising avenue of further work is to use these expressions to empirically evaluate the task-based model predictions for heterogeneity in robot adoption treatment effects.

C Structural Estimation of Firm Robot Adoption

C.1 Elasticity of Substitution Between Production Tasks

This section uses the model presented in Section 3 to derive the moment condition that I use to estimate the elasticity of substitution between production tasks σ in Section 4.1. The derivations follow closely those in Doraszelski and Jaumandreu (2018).

To derive the moment conditions, first insert Equation (13) into Equation (9) to express the deterministic component of firm productivities in terms of a non-parametric function of observables

$$\varphi_{ojt} = g_{ot}(\varphi_{ojt-1}, ..., \varphi_{ojt-k}) + \xi_{ojt} \tag{47}$$

$$= g_{ot}\left(l_{ojt-1} - m_{jt-1} + \sigma(w_{ojt-1} - w_{Mjt-1}), .., l_{ojt-k} - m_{jt-k} + \sigma(w_{ojt-k} - w_{Mjt-k})\right) + \xi_{ojt} \tag{48}$$

$$= h_{ot}(l_{ojt-1} - m_{jt-1}, w_{ojt-1} - w_{Mjt-1}, .., l_{ojt-k} - m_{jt-k}, w_{ojt-k} - w_{Mjt-k}) + \xi_{ojt}, \tag{49}$$

where lower-case letters denote log-transforms. Insert this function into Equation (11) to obtain

$$l_{o'jt} - l_{ojt} = -\sigma(w_{o'jt} - w_{ojt}) + (h_{o'jt} - h_{ojt}) + (\xi_{o'jt} - \xi_{ojt}), \tag{50}$$

Equation (50) holds for firms that have not yet adopted robots. The Markovian structure on firm productivities implies that past factor choices l_{jt} and prices w_{jt} have to be uncorrelated with the current productivity innovations ξ_{jt} that constitute the error term in Equation (50). I can thus form a population moment condition that identifies σ, my

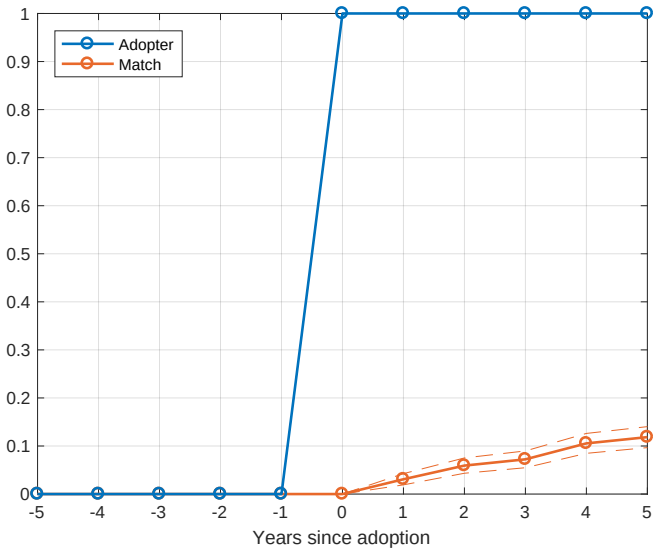
parameter of interest

$$\mathbb{E}_t \left[A_{oo'}(Q_{jt-1}) \left(l_{o'jt} - l_{ojt} + \sigma(w_{o'jt} - w_{ojt}) - (h_{o't} - h_{ot}) \right) \right] = 0, \tag{51}$$

where $A_{oo'}$ is a vector function of the instruments Q_{t-1} including l_{jt-1}, w_{jt-1}. The instrument vector Q_t consists of quadratic functions of $l_{jt-k} - m_{jt-k}$ and $w_{t-k} - w_{Mt-k}$ for $k = 1, 2, 3$, as well as quadratic functions of w_{jt-1} and l_{jt-1} (the excluded instruments). I set "Production Workers" and "Tech Workers" as o and o', respectively, and I use "Other Workers" as the benchmark factor in production (M in the derivations above). I estimate (51) using a two-step GMM procedure (the `gmm` package in Stata).

C.2 Robot Technology

Figure C.1: Firm Robot Adoption Around the Event Year



Notes: The figure shows separately the shares of firms in the treatment and control groups that have adopted robots around the event year.

C.2.1 Econometric Specification of the Event Studies

In this section, I describe the econometric specification that generates the matching-based event study estimates plotted in Figures 1 and 2. The estimates are differences-in-differences of outcomes y_{jt} for robot adopters versus match firms measured relative to the year prior to adoption.[30] Figures 1 and 2 plot OLS estimates of δ_k from the following specification

$$\frac{y_{jt}}{y^e_{j\text{pre}}} = \alpha \times \mathbb{R}_{je} + \sum_{k \in \mathcal{K}} \alpha_k \times \mathbb{1}_{\{t=e+k\}} + \sum_{k \in \mathcal{K} \setminus \{-1\}} \delta_k \times \mathbb{1}_{\{t=e+k\}} \times \mathbb{R}_{je} + u_{jt} \tag{52}$$

where e denotes event year, $y^e_{j\text{pre}}$ are pre-event median outcomes, $\mathbb{R}_{je}$ indicates that firm j adopted robots in year e, and $\mathbb{1}_{\{t=e+k\}}$ is an indicator that switches on iff event year e occurred k years ago. The event study window is denoted $\mathcal{K} = [-4, 4]$. Standard errors are clustered at the match level. I allow for zeros in occuptional wage bills by calculating this relative difference as $(y_{jt}/w_{jpre})/(y_{pre}/w_{pre})$, where w_{jt} is the total wage bill of firm j in year t, and y_{pre} denotes mean pre-event outcomes.

C.2.2 Robot Technology Distributed Lag Model

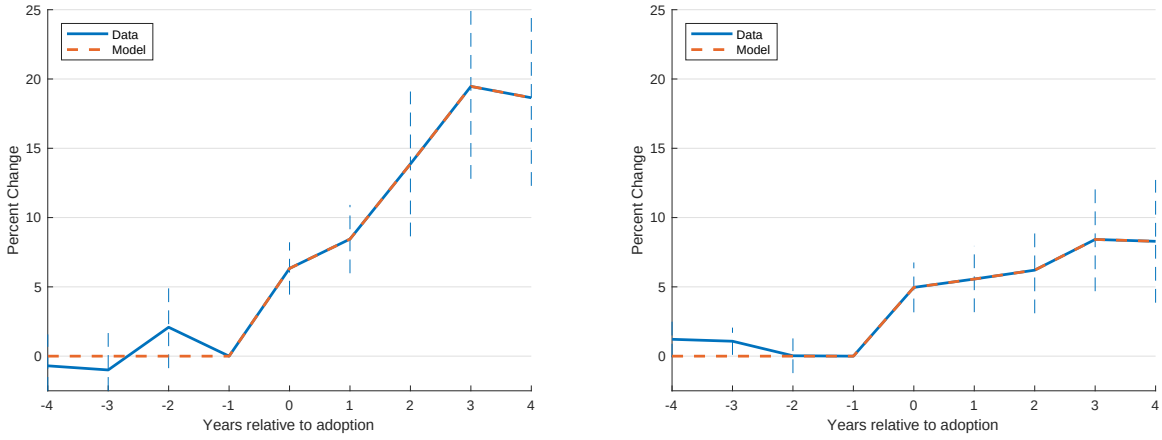
This section generalizes the robot technology equations (2)-(3) to account for the dynamic adjustments to robot production observed in Figures 1 and 2. I let robot technology follow a distributed lag model

$$\log(z_{jt}) = \varphi_{jt} + \sum_{\tau=0}^{4} \gamma_\tau R_{jt-\tau} \tag{53}$$

[30] The match firms are found using an Exact-Mahalanobis matching procedure described in Appendix A.7.1.

Following the identification argument in Section 4.2.1, the adoption event study moments in Figures 1 and 2 exactly identify the dynamic robot technology parameters γ_τ. Figure C.2 shows the model fit for firm sales and wage bills.

Figure C.2: Distributed Lag Model for Robot Productivities



(a) Sales

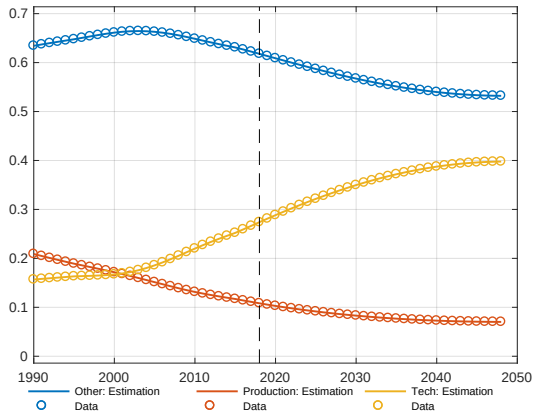
(b) Wage Bill

C.3 Baseline Technology

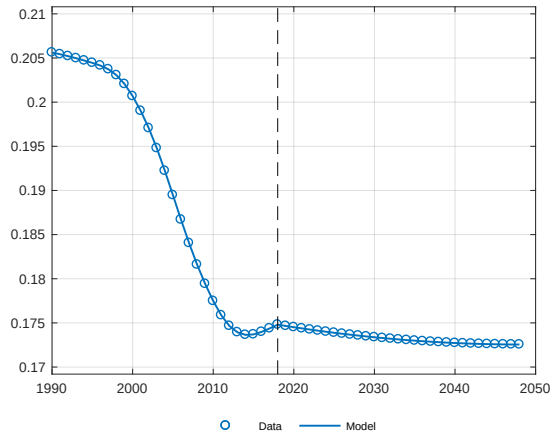
C.3.1 Labor-Augmenting Baseline Productivities

The general equilibrium model restricts the labor-augmenting baseline productivities to a common time-varying parameter vector. I calibrate this path of common productivities γ_{ot} to match the aggregate factor shares in manufacturing taking into account the diffusion of robot technology. Figure C.3 shows data (dots) and model simulations (line) from 1990 to 2018 together with out-of-sample forecasts from 2019 to 2049. The data have been HP-filtered to focus on medium-run movements (smoothing parameter of 100 following Backus et al. (1992)). The forecasts extrapolate the growth rate from 2011 to 2018, assuming a linear reduction in rates of growth to zero by 2049.

Figure C.3: Aggregate Factor Shares in Manufacturing Production



(a) Wage Bill Shares



(b) Wage Bill Share in Manufacturing Sales

C.3.2 Hicks-Neutral Baseline Productivities

With the homogeneity restriction imposed on firm baseline labor-augmenting productivities, the productivity process in Equation (21) boils down to an AR(1) process for the Hicks-neutral term φ_{Hjt}.

$$\varphi_{Hjt} = \mu_{Ht} + \rho_H \varphi_{Hjt-1} + \sigma_H \epsilon_{jt}, \tag{54}$$

where ρ_z is the persistence parameter for baseline productivity, and ψ_t is a time fixed effect.

Table C.1: Baseline Productivity Parameters

Parameter	Description	Estimated Value
$\hat{\rho}_H$	Persistence of firm productivity	0.901 (0.062)
$\hat{\sigma}_H$	Standard deviation of firm productivity innovations	0.140

I discretize the estimated AR(1) process using the Tauchen (1986) method.

C.4 Robot Adoption Costs

C.4.1 Method of Simulated Moments (MSM) Estimator

In this section, I describe the method of simulated moments (MSM) estimation procedure adopted in Section 4.4. Table C.2 reports the MSM parameter estimates.

1. Parameterize robot adoption costs to be log-linear in time: $c_t^R = \exp(c_0^R + c_1^R \times t)$.

2. Stack the robot adoption cost parameters into the parameter vector $\theta = (c_0^R, c_1^R, \nu)'$.

3. Stack the robot diffusion curve and the adopter size premium into the moment vector $\pi \in \mathbb{R}^N$ with $N = 2018 - 1990 + 2$

4. Define a grid on the parameter space Θ. For each point on the grid $\theta^{(j)} \in \Theta$,

   (a) Solve for continuation values given $c_t^R = \exp(c_0^{(j)} + c_1^{(j)} \times t)$ and $\nu = \nu^{(j)}$. The solution algorithm is specified in Section F.1.

   (b) Simulate firms forward using policy functions.

   (c) Calculate the in-sample squared deviations between the simulated and observed moment vectors

   $$(\pi_S(\theta^{(j)}) - \pi_D)' W (\pi_S(\theta^{(j)}) - \pi_D) \tag{55}$$

   where W is the identity weighting matrix.

5. The MSM estimator, $\hat{\theta}$, attains the minimum in (55).

Table C.2: Robot Adoption Cost Parameters (MSM)

Parameter	Description	Estimate
c_0^R	Intercept of the common adoption cost schedule over time	1.155
c_1^R	Slope of the common adoption cost schedule over time	-0.026
ν	Dispersion in idiosyncratic adoption costs	0.384

C.4.2 Variance-Covariance Matrix of the MSM Estimator

I calculate the variance-covariance matrix of the MSM estimator $\hat{\theta}$ using the formula on pages 88 and 89 of Adda and Cooper (2003). The MSM estimator has the asymptotic distribution

$$\sqrt{N}(\theta - \theta_0) \xrightarrow{d} \mathcal{N}(0, V) \quad \text{with} \tag{56}$$

$$V = \left[\mathbb{E}_0 \frac{\partial \pi'}{\partial \theta} W^{-1} \frac{\partial \pi}{\partial \theta'}\right]^{-1} \mathbb{E}_0 \frac{\partial \pi'}{\partial \theta} W^{-1} \Sigma(\theta_0) W^{-1} \frac{\partial \pi}{\partial \theta'} \times \left[\mathbb{E}_0 \frac{\partial \pi'}{\partial \theta} W^{-1} \frac{\partial \pi}{\partial \theta'}\right]^{-1}, \tag{57}$$

I estimate $\mathbb{E}_0 \frac{\partial \mu'}{\partial \theta}$ using numerical derivatives of the simulated moments around $\hat{\theta}$. The confidence bands in Figure 3 are calculated using the delta method.

C.4.3 Comparison of Robot Adoption Cost Estimates

Table C.3 compares the MSM estimate for the rate of change in the common component of robot adoption costs c_1^R to the robot machine expenditures reported on customs forms of adopting firms. I report the time-slope estimates of log expenditures and log mean expenditures. As the table shows, the MSM estimate of c_1^R falls within the confidence bands of both specifications.

Table C.3: Rate of Change in Robot Adoption Costs: Model Estimates vs. External Measures

	Lower Bound (95% CI)	Point Estimate	Upper Bound (95% CI)
MSM Estimate ($\hat{c}_1^R$)		-0.0264	
Customs Expenditures 1	-0.1158	-0.0693	-0.0229
Customs Expenditures 2	-0.0661	0.0179	0.1019

Notes: The first row is the MSM estimate of $\hat{c}_1^R$. The second row (Customs Expenditures 1) is the OLS estimate of β_1 in $log(Y_{jt}) = \beta_0 + \beta_1 t$ (reweighted to the yearly level). The third row (Customs Expenditures 2) is the OLS estimate of β_1 in $\log(\bar{Y}_t) = \beta_0 + \beta_1 t$. I deflate the customs expenditures with the consumer price index.

C.5 Depreciation of Robot Technology

This section derives a model extension in which robot technology deteriorates with a probability θ. The Bellman equation for robot adoption now reads

$$V_t(0, \varphi) = \max_{R \in \{0,1\}} \pi_t(0, \varphi) - (c_t^R + \varepsilon_{jt}^R) \times R + \beta \mathbb{E}_t V_{t+1}(R, \varphi') \tag{58}$$

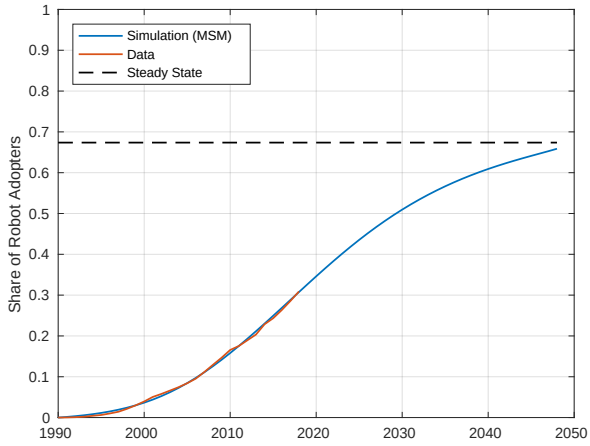
$$V_t(1, \varphi) = \pi_t(1, \varphi) + (1 - \theta)\mathbb{E}_t V_{t+1}(1, \varphi') + \theta \mathbb{E}_t V_{t+1}(0, \varphi') \tag{59}$$

Equations (58)-(59) collapse to the current setup in Equations (7)-(8) if $\theta = 0$. Figure C.4 shows the simulated robot diffusion curve and real wage effects on industrial robots under a robot depreciation rate θ of 10 percent (the depreciation rate used in Graetz and Michaels (2018)). Compared to baseline Figures 3a and 6, the model extension to robot depreciation does not affect the in-sample estimate of the real wage effects of industrial robots as the extended model is estimated to match the same observed robot diffusion curve. The model extension does alter the long-run predictions, however, as the robot

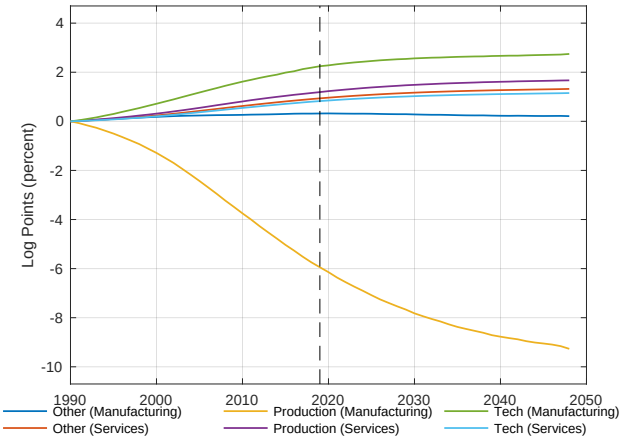
diffusion curve asymptotes to a long-run steady-state level below full adoption when robot technology deterioates (dashed line in Figure C.4a).

Figure C.4: Effect of Industrial Robots with Depreciation of Robot Technology

(a) Robot Diffusion Curve



(b) Real Wage Effects



D The Labor Supply Block

D.1 Data on Worker Transitions

This section describes how I measure the worker transitions used to estimate the labor supply model. I follow Traiberman (2019) as closely as possible; please refer to his Appendices B.5 and D for additional details.

I use the Integrated Database for Labor Market Research (IDA), which links every worker to her employer in the month of November. The data contain information about the occupation, salary, and sociodemographics of workers. I follow the recommendation of Statistics Denmark, and use only the high-quality occupational codes that come from administrative registers or pension funds (`DISCOTYP` 1, 2, 4 and 10). If a worker is em-

ployed in the same occupation in year $t-1$ and $t+1$ but has missing data at t, I impute the occupation as that at $t-1$. If a worker switches firms between year t and $t+1$ and there are data on occupation in $t+2$, I impute any missing occupational codes at $t+1$. As the measure of salary $w_{ot}H_{oit}$, I use the variable `joblon` together with a correction for labor market pension contributions from Hummels et al. (2014). The fine level of disaggregation into worker types and occupations implies that the observed transition matrix $\pi_t(s, a, \text{ten}, o, o')$ has zero elements. To avoid bias from dropping these zero elements, I follow Traiberman (2019) and perform the following first-stage smoothing regression of the transition probabilities.

$$\pi_t(s, a, \text{ten}, o, o') = \beta^0_{soo't} + \beta^a_{soo't}a + \beta^{a2}_{soo't}a^2 + \beta^{a3}_{soo't}a^3 \tag{60}$$

$$+ \beta^{\text{ten}}_{soo't}\text{ten} + \beta^{\text{ten2}}_{soo't}\text{ten}^2 + \beta^{\text{ten3}}_{soo't}\text{ten}^3 + \beta^{\text{ten},a}_{soo't}a \times \text{ten} \tag{61}$$

I then use the predicted transition rates $\hat{\pi}_t$ (bounded by 10^{-6} and $1-10^{-6}$) as input variable into the estimating regression equations (31)-(33).

D.2 Estimation of the Labor Supply Model

Table D.1: Human Capital Function

	Tech (services)	Tech (manuf)	Production (services)	Production (manuf)	Other (services)	Other (manuf)
Age β_1^o	0.0285	0.0265	0.0096	0.0055	0.0124	0.0139
	(0.0010)	(0.0005)	(0.0006)	(0.0006)	(0.0007)	(0.0010)
Age-Squared β_2^o	-0.0590	-0.0543	-0.0236	-0.0171	-0.0266	-0.0301
	(0.0016)	(0.0013)	(0.0011)	(0.0014)	(0.0014)	(0.0023)
Tenure β_3^o	0.0300	0.0153	0.0277	0.0234	0.0537	0.0307
	(0.0018)	(0.0010)	(0.0012)	(0.0016)	(0.0030)	(0.0012)
Mid Skill β_M^o	-0.0428	0.0028	0.1025	0.1168	0.0537	0.1165
	(0.0015)	(0.0028)	(0.0015)	(0.0025)	(0.0012)	(0.0018)
High Skill β_H^o	0.1671	0.2958	0.0997	0.1629	0.2502	0.5108
	(0.0016)	(0.0022)	(0.0103)	(0.0061)	(0.0037)	(0.0053)
Observations	2147314	602741	1029836	681133	17176380	2780515

Notes: SD of income shock: Tech (services): .118, Tech (manufacturing): .077, Production (services): .096, Production (manufacturing): .077 Others (services): .148, Others (services): .133. Standard errors are clustered at the occupation-year level. Coefficient on Age Squared is presented $\times 10^2$.

D.2.1 Occupational Switching Costs

The non-linear least squares objective function (NLLS) reads

$$\min_{\{c,\rho\}} \sum_{\omega,o,o',t} \left[\log \frac{\pi_t(oo'|\omega)}{\pi_t(oo|\omega)} + \beta \log \frac{\pi_{t+1}(o'o''|\omega')}{\pi_{t+1}(oo''|\omega'')} - \left(\frac{1}{\rho} c_{oo'}(\omega) - \frac{\beta}{\rho}(c_{o'o''}(\omega') - c_{oo''}(\omega'')) + \frac{\beta}{\rho}(w_{o't+1} H_{o'}(\omega') - w_{ot+1} H_o(\omega'')) + \frac{\beta}{\rho}(\eta_{o'} - \eta_o) \right) \right]^2 \tag{62}$$

I use the Matlab solver `lsqnonlin` to estimate $c(\omega)$ and ρ in Equation (62). Table D.2 presents the estimated bilateral occupational switching costs, and Table D.3 presents the remaining switching cost estimates.

Table D.2: Bilateral Switching Costs $c_{oo'}/\rho$

	Tech (serv)	Tech (manuf)	Production (serv)	Production (manuf)	Other (serv)	Other (manuf)
Tech (services)	0	4.71	1.7	5.04	1.17	4.64
Tech (manufacturing)	0.76	0	3.49	0.58	2.51	0.01
Production (services)	5.9	11.12	0	2.73	3.78	6.6
Production (manufacturing)	9.24	8.79	2.75	0	6.35	4.28
Other (services)	3.8	8.44	1.87	3.9	0	2.97
Other (manufacturing)	6.6	5.94	3.68	1.27	2.19	0

Table D.3: Switching Cost Parameters

Parameter	Description	Estimate
α_1	Semi-elasticity of switching costs with respect to age (linear term)‡	13.86
α_2	Semi-elasticity of switching costs with respect to age (quadratic term)‡	-0.14
α_M	Semi-elasticity of switching cost with respect to mid skill	0.01
α_H	Semi-elasticity of switching cost with respect to high skill	0.00
ρ	Occupational preference shock variance†	2.00

Notes: ‡ Coefficients of age polynomial are presented $\times 10^3$. †Parameter value of ρ used in Section 6.

D.2.2 Occupational Amenities

The employment shares across occupations are tightly connected to the vector of occupational amenities, η_{ot}. Conditional on the distribution of workers in $t-1$, the relative change in employment shares, $\hat{s}_{ot} = \frac{s'_{ot}}{s_{ot}}$ from a different occupational amenity η'_{ot} is given by

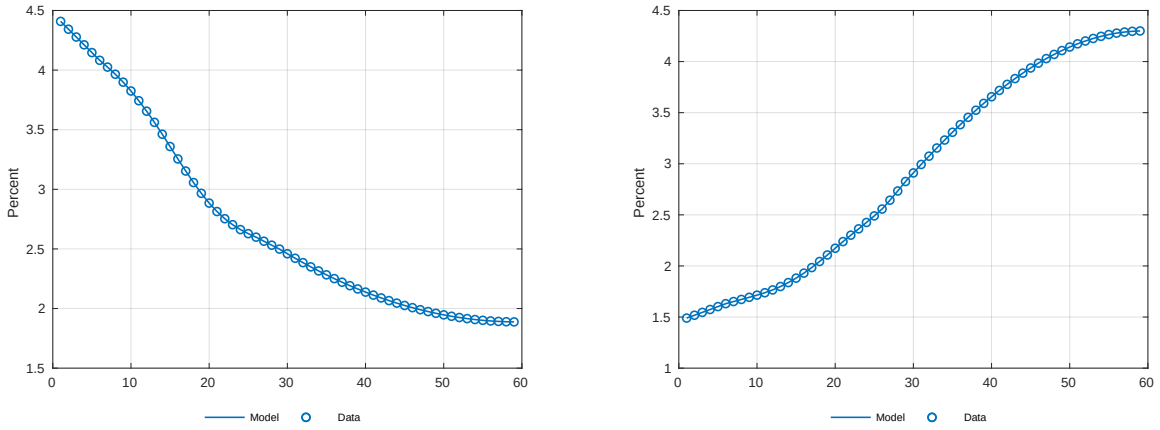
$$\hat{s}_{o't} = \frac{\exp(\frac{1}{\rho}\hat{\eta}_{o't})}{\sum_{o''} s_{o''t} \times \exp(\frac{1}{\rho}\hat{\eta}_{o''t})}. \tag{63}$$

Of course, when evaluating the total effect of η_{ot} on s_{ot}, I cannot take the distribution of worker states, μ^W_{t-1}, as given because workers also choose occupations before $t-1$ in

anticipation of the amenities in year t. In practice, I estimate the path of occupational amenities η_{ot} by matching simulated occupational employment shares to the data. The estimation procedure is a fixed-point shooting algorithm.

Figure D.1 shows data (dots) and model simulations (line) for the share of employment across two example occupations from 1990 to 2018 together with out of sample forecasts from 2019 to 2049. The data have been HP-filtered to focus on medium-run movements (smoothing parameter of 100 following Backus et al. (1992)). The forecasts extrapolate growth rates from 2011 to 2018 by assuming a linear reduction in rates of growth to zero by 2049.

Figure D.1: Employment Shares Across Occupations (Manufacturing)



(a) Production Workers

(b) Tech Workers

E Model Extension to Firm-Specific Wages

This section proposes an extension of the model in Sections 3 and 5 that accommodates firm-specific wages in equilibrium. On the worker side, I embed a random utility model for workplace environments as in Card et al. (2018) into the dynamic discrete occupational

choice model of Section 5. The heterogeneity in worker preferences for employers implies that firms face upward-sloping labor supply functions. I derive the cost, profit, and factor demand functions of the firm, and I clarify how to implement these expressions into the remaining model structure.

E.1 The Worker's Problem

The worker first chooses which occupation o' to work in next period. Upon arriving in the chosen occupation next period, the worker then chooses which employer $j \in \mathcal{J}_{o'}$ to work for. This job choice model is a nested logit with occupations in the upper nest and firms constituting the lower nest. The Bellman equation of the worker reads

$$v_t(o,s,\text{age},\text{ten}) = \max_{j \in \mathcal{J}_o} \left\{ \log(w_{ojt} H_o(s,\text{age},\text{ten})) + \log(a_{ojt}) + \varepsilon_{oijt} \right\} \tag{64}$$

$$+ \max_{o' \in \mathcal{O}} \left\{ -(c_{oo'} + \varepsilon_{o'}) + \mathbf{1}_{\{\text{age}<65\}} \beta \mathbb{E}_t v_{t+1} \left(o', s, \text{age}+1, \mathbf{1}_{\{o'=o\}} (\text{ten}+1) \right) \right\} \tag{65}$$

where a_{ojt} is a firm-specific amenity common to all workers, and $\varepsilon_{oijt} \overset{iid}{\sim} \text{GEV1}(\alpha)$ are idiosyncratic workplace preference shocks that are drawn from a Gumbel distribution and only realized once a worker arrives in the occupational labor market. The parameter α measures the dispersion in these idiosyncratic preference shocks. The probability that a worker in occupation o with characteristics ω chooses to work for firm j is

$$P\left(j_t(\omega,o) = j | \omega\right) = \frac{(a_{ojt} w_{ojt} H_o(\omega))^{1/\alpha}}{\sum_{j' \in \mathcal{J}_o} (a_{oj't} w_{oj't} H_o(\omega))^{1/\alpha}} \tag{66}$$

E.2 The Firm's Problem

The firm faces the labor supply curve

$$L_{ojt}(w_{ojt}) = \int \frac{(a_{ojt}w_{ojt}H_o(\omega))^{1/\alpha}}{\sum_{j' \in \mathcal{J}_o}(a_{oj't}w_{oj't}H_o(\omega))^{1/\alpha}} H_o(\omega) dF_t^W(\omega) \tag{67}$$

If individual firms each constitute a negligible share of the total occupational labor market, the inverse labor supply curve to the firm becomes

$$w_{ojt} = \frac{w_{ot}}{a_{ojt}} \times L_{ojt}^{\alpha} \tag{68}$$

with $w_{ot} = \{\int \frac{H_o(\omega)^{1/\alpha}}{\sum_{j' \in \mathcal{J}_o}(a_{oj't}w_{oj't}H_o(\omega))^{1/\alpha}} dF_t^W(\omega)\}^{-\alpha}$.[31] To ease the exposition, stack the static inputs (including intermediate inputs) into the vector L, and reparameterize, without loss of generality, the CES production function as follows

$$Y_{jt} = F(L_{jt}|z_{jt}) = \left\{ \sum_o z_{ojt}^{\frac{1}{\sigma}} L_{ojt}^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} \tag{69}$$

The firm's profit maximization problem reads

$$\pi_t(z_{jt}) = \max_{L_{jt}} \left\{ P_{Mt} Y_{Mt}^{\frac{1}{\epsilon}} F(L_{jt}|z_{jt})^{1-1/\epsilon} - \sum_o \frac{w_{ot}}{a_{ojt}} L_{ojt}^{1+\alpha} \right\}, \tag{70}$$

[31] If firms are price takers in intermediate input (M) markets, then the specification should be amended to allow for $\alpha \neq \alpha_M = 0$.

The first-order conditions for cost minimization imply that

$$\log(L_{o'jt}) - \log(L_{ojt}) = -\frac{\sigma}{1+\alpha\sigma}(\log(w_{o'jt}) - \log(w_{ojt})) \tag{71}$$

$$+ \frac{1}{1+\alpha\sigma}\{\sigma \log(a_{ojt}/a_{o'jt}) + \log(z_{o'jt}/z_{ojt})\}, \tag{72}$$

which clarifies that σ in Equation (11) is estimated off the variation in firms' relative amenities $\frac{a_{ojt}}{a_{o'jt}}$ once I in Section 4.1 control for the evolution in firm productivities $\frac{z_{ojt}}{z_{o'jt}}$. Equation (71) shows that the GMM estimate $\hat{\sigma}$ is attenuated toward zero if workers have heterogeneous preferences for workplaces ($\alpha > 0$).

The cost function of the firm is (after some algebra)

$$C_t(Y_{jt}, z_{jt}) = \min_{L_{jt}} \sum_o w_{ojt}(L_{ojt})L_{ojt} \quad \text{s/t} \quad F(L_{jt}|z_{jt}) \geq Y_{jt} \tag{73}$$

$$= Y_{jt}^{1+\alpha}\left\{\sum_o (z_{ojt}^{1+\alpha}(\frac{w_{ot}}{a_{ojt}})^{1-\sigma})^{\frac{1}{1+\alpha\sigma}}\right\}^{\frac{1}{1-\sigma}} = Y^{1+\alpha}\Omega_{jt}, \tag{74}$$

where I denote $\Omega_{jt} = \{\sum_o (z_{ojt}^{1+\alpha}(\frac{w_{ot}}{a_{ojt}})^{1-\sigma})^{\frac{1}{1+\alpha\sigma}}\}^{\frac{1}{1-\sigma}}$. The profit-maximizing output level is

$$Y_{jt} = \left\{Y_{Mt}P_{Mt}^{\epsilon}((1+\alpha)\frac{\epsilon}{(\epsilon-1)}\Omega_{jt})^{-\epsilon}\right\}^{\frac{1}{1+\alpha\epsilon}} \tag{75}$$

The firm's factor demands are

$$L_{ojt} = \frac{\{z_{ojt}(\frac{w_{ot}}{a_{ojt}})^{-\sigma}\}^{\frac{1}{1+\alpha\sigma}}}{\Omega_{jt}^{-\sigma}} \times Y_{jt} \tag{76}$$

The profit function is

$$\pi_t(z_{jt}) = \left(\frac{1-\alpha\epsilon}{\epsilon-1}\right)\left\{\Omega_{jt}^{1-\epsilon(1-\alpha)}Y_{Mt}P_{Mt}^{\epsilon}\left((1+\alpha)\frac{\epsilon}{\epsilon-1}\right)^{-\epsilon}\right\}^{\frac{1}{1+\alpha\epsilon}} \tag{77}$$

Equations (71)-(77) simplify to the usual CES expressions when worker preferences for workplaces vanish, $\alpha \to 0$.

The flow profit function (77) can be directly plugged into the Bellman equation (7) of the adoption model in Section 3. The firm factor demands (76) can be readily aggregated up when clearing labor markets in the general equilibrium model of Section 6.1.

F Solution Algorithms

This section provides details on the solution algorithms used in Sections 4, 5, and 6.

F.1 Solving the Firm's Problem

This section details the algorithm for solving the firm's dynamic programming problem of robot adoption.

1. Set a time horizon, T, sufficiently far in the future such that robots are fully diffused and robot adoption costs are stationary by then (I set $T = 2050$ in practice).
2. Start at T. Solve the stationary, infinite horizon dynamic programming problem by iterating on the expected value functions until convergence.

$$\mathbb{E}V_T^{(j+1)}(1,\varphi) = \pi_T(1,\varphi) + \beta\sum_{z'} p(\varphi'|\varphi)\mathbb{E}V_T^{(j)}(1,\varphi') \tag{78}$$

$$\mathbb{E}V_T^{(j+1)}(0,\varphi) = \pi_T(0,\varphi) + \beta\sum_{z'} p(\varphi'|\varphi)\nu\log\left\{\exp(\frac{1}{\nu}(-c_T^R + \beta\mathbb{E}V_T^{(j)}(1,\varphi'))) + \exp(\frac{1}{\nu}\beta\mathbb{E}V_T^{(j)}(0,\varphi'))\right\}, \tag{79}$$

where I used the log-sum expression for the expected maximum (EMAX) function.[32] Convergence of Equation (79) in the unique fixed point $\mathbb{E}V_T(R, \varphi)$ is ensured from Blackwell's sufficient conditions for contraction mappings (Stokey and Lucas, 1989, Theorem 4.6).

3. Solve for $\{\mathbb{E}V_t(R, \varphi)\}_{t=t_0}^{T-1}$ using backward recursion from $T-1$ to the initial period t_0.

$$\mathbb{E}V_t(1, \varphi) = \pi_t(1, \varphi) + \beta \sum_{z'} p(\varphi'|\varphi) V_{t+1}(1, \varphi') \tag{80}$$

$$\mathbb{E}V_t(0, \varphi) = \pi_t(0, \varphi) + \beta \sum_{z'} p(\varphi'|\varphi) \nu \log \left\{ \exp(\frac{1}{\nu}(-c_t^R + \beta \mathbb{E}V_{t+1}(1, \varphi'))) + \exp(\frac{1}{\nu}\beta \mathbb{E}V_{t+1}(0, \varphi')) \right\} \tag{81}$$

4. From the initial year t_0, use policy functions to simulate firms forward. Verify that the robot adoption share is 1 at time T.

In solving steps 3 and 4, I assume that firms have perfect foresight with respect to the in-sample path of wages, and use regressions to forecast these aggregate state variables out of sample.

F.2 Solving the Worker's Problem

This section details the algorithm for solving the worker's dynamic occupational choice problem.

1. Set a time horizon, T, sufficiently far in the future such that robots are fully diffused by then (I set $T = 2050$ in practice).

[32] Note that my setup with a logit shock for adoption (Equation (22)) is isomorphic to the setup in Rust (1987) with Gumbel shocks for both adoption and non-adoption (up to a recentering for the mean of a Gumbel). This is due to the well-known result that the difference between two Gumbels is logistically distributed.

2. Start at T. Solve the stationary worker value functions

   (a) Start at age of retirement. The value function is

$$\mathbb{E}_{\varepsilon,\zeta} V_T(o, 65, \omega) = \log(w_{oT} H_{oT}(65, \omega)) + a_{oT}. \tag{82}$$

   (b) Solve the value function for ages $a = 64, .., 25$ by backward recursion

$$\mathbb{E}_{\varepsilon,\zeta} V_T(o, a, \omega) = \log(w_{oT} H_{oT}(a, \omega)) + a_{oT} + \rho \left[\gamma + \log \left\{ \sum_{o'} \exp(\frac{1}{\rho}(c_{oo'}(\omega) + \beta \mathbb{E}_{\varepsilon,\zeta} V_T(o', a+1, \omega'))) \right\} \right], \tag{83}$$

   where $\gamma = 0.577$ is Euler's constant.

3. Calculate the worker value functions for $t = T-1, ..., t_0$ using backward recursion

$$\mathbb{E}_{\varepsilon,\zeta} V_t(o, 65, \omega, \zeta) = \log(w_{ot} H_{ot}(65, \omega)) + a_{ot} \tag{84}$$

$$\mathbb{E}_{\varepsilon,\zeta} V_t(o, a, \omega, \zeta) = \log(w_{ot} H_{ot}(a, \omega)) + a_{ot} + \rho \left[\gamma + \log \left\{ \sum_{o'} \exp(\frac{1}{\rho}(c_{oo'}(\omega) + \beta \mathbb{E}_{\varepsilon,\zeta} V_{t+1}(o', a+1, \omega'))) \right\} \right], \tag{85}$$

In solving this dynamic program, I assume that workers have perfect foresight with respect to the in-sample path of wages, and use regressions to forecast these aggregate state variables out-of-sample.

F.3 Solving the Dynamic General Equilibrium

This section describes the algorithm for solving the general equilibrium featuring the two-sided dynamics defined in Section 6.1. A key property of the general equilibrium model is that, despite the rich worker and firm heterogeneity, the only aggregate state variables that agents need to keep track of to solve their dynamic programming problem is the path

of wages and the manufacturing price index.[33] I use a fixed-point shooting algorithm that solves for the wage path that clears labor markets given the optimal policy functions of workers and firms.

1. Guess a path of wages $w_t^{(0)}$ and manufacturing price index P_{Mt}^0.

2. Solve for firm and worker continuation values (see Appendices F.1 and F.2).

3. Simulate firm and worker states forward using the policy functions from Step 2.

4. Find wages, $w_t^{(e)}$, that clear labor markets for each occupation period by period (using the firms' static labor demand conditions from Equation (5)). Calculate the implied manufacturing price index $P_{Mt}^{(e)}$.

5. Update wages and manufacturing price index

$$w_t^{(j+1)} = \lambda w_t^{(j)} + (1-\lambda) w_t^{(e)} \tag{86}$$

$$P_{Mt}^{(j+1)} = \lambda P_{Mt}^{(j)} + (1-\lambda) P_{Mt}^{(e)} \tag{87}$$

   where $\lambda \in [0.8, 0.95]$ is the relaxation parameter in the Gauss-Seidel update.

6. Iterate until convergence in $\{w_t, P_{Mt}\}_t$.

[33] The path of wages are sufficient to solve the worker's problem. Manufacturing firms also need to keep track of the manufacturing output price index as it summarizes the competitive pressures from robot adoption.

G Counterfactual Experiments

Table G.1: Parameters of the General Equilibrium Model

	Description	Related Moments	Time varying
Manufacturing Firms			
c_t^R	Common robot adoption costs	Robot diffusion curve (Figure 3)	✓
ν	Variance of idiosyncratic adoption costs	Size premium in robot adoption (Figure 4)	
γ_o	Labor-augmenting robot productivity	Robot adoption event studies (Figures 1-2)	
γ_H	Hicks-neutral robot productivity†	Robot adoption event studies (Figures 1-2)	
σ	Elasticity of task substitution	Rational expectations GMM (Table 3)	
μ_H	Mean of Hicks-neutral baseline productivity	Real wage index	✓
ρ_H	Persistence of Hicks-Neutral productivity	Firm sales dynamics (Table C.1)	
σ_H	Standard deviation of Hicks-Neutral innovations	Firm sales dynamics (Table C.1)	
φ_{ot}	Baseline labor-augmenting producitivites	Labor shares in manufacturing sales (Figure C.3)	✓
Workers			
β	Human capital parameters	Mincer regression (Table D.1)	
$c_{oo'}$	Occupational switching costs	Occupational transition rates (Tables D.2-D.3)	
η_{ot}	Occupational amenities	Employment shares across occupations and sectors (Figure D.1)	✓
Services Production			
α^s	Cobb Douglas shares in services production	Wage bill shares in sales excl. manufacturing	
z_{st}	Hicks-Neutral productivity in services	Real wage index	✓
Common Parameters			
β	Discount factor	Interest rate of 4%	
μ	Cobb-Douglas shares in final output	Share of manufacturing in total output	
ϵ	Elasticity of manufacturing demand	Markup of 1/3 (Bloom, 2009)	

Notes: †I calibrate the path of γ_{Ht} to hold the sales elasticity with respect to robot adoption (Figure 1(a)) constant over time, given the estimated non-stationarity path of baseline labor-augmenting productivites γ_{ot}.

G.1 The Distributional Impacts of Industrial Robots

Figure G.1: The Effect of Industrial Robots on the Labor Share in Manufacturing Sales

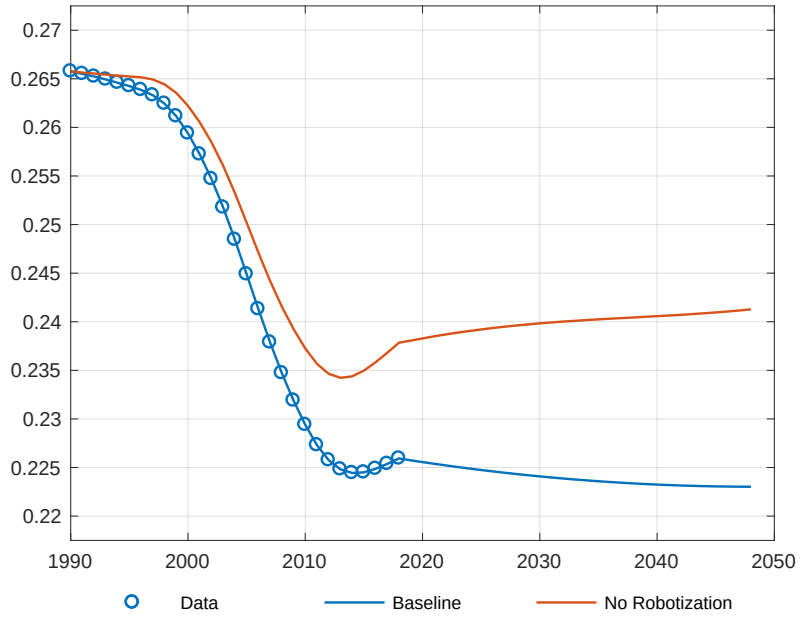
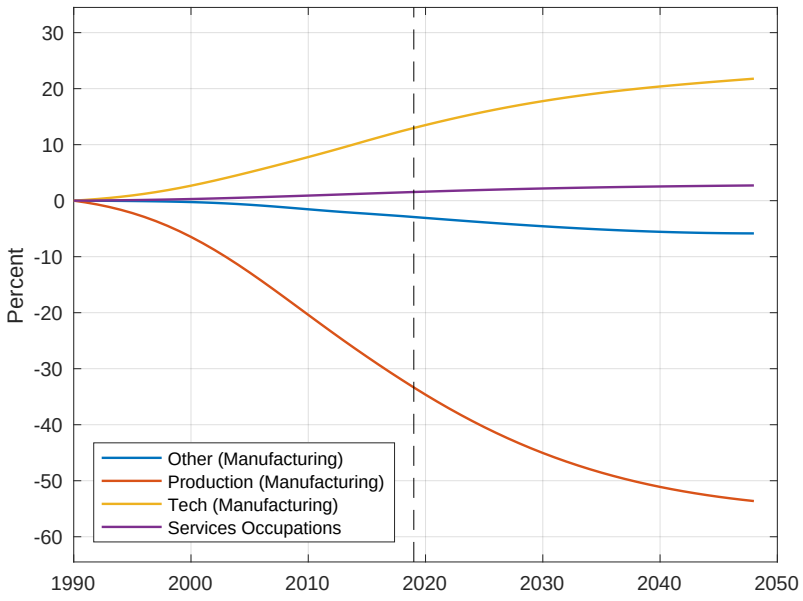


Figure G.2: Real Wage Effects of Industrial Robots with Exogenous Labor Supply



G.1.1 Compensating Variations

To measure the effects on worker welfare, I follow Caliendo et al. (2019) and calculate the percentage annual wage change δ needed to compensate a worker of characteristics ω and age a for a given change in policy. Let v^0 and v^1 denote the worker value functions in two policy scenarios whose welfare implications we would like to compare. Due to the logarithmic flow utility of workers in Equation (27), the compensating variations δ are simply given by

$$v_t^1(\omega, a) = v_t^0(\omega, a) + \sum_{\tau=0}^{\bar{A}-a} \beta^\tau \delta_t(\omega, a) \tag{88}$$

$$\delta_t(\omega, a) = (v_t^1(\omega, a) - v_t^0(\omega, a)) \frac{(1-\beta)}{(1-\beta^{\bar{A}-a+1})} \tag{89}$$

G.2 Policy Counterfactual: The Dynamic Incidence of a Robot Tax

G.2.1 Predatory Investment Externalities

This section incorporates predatory investment effects into the robot tax incidence analysis. Predatory investment effects refer to the pecuniary externality where parts of the profit gains from robot adoption come from crowding out competitors in output markets. If demand is sufficiently elastic, firms will be willing to undertake very costly fixed robot investments to obtain just an infinitesimal variable cost advantage over its competitors.

To analyze the effects of such predations, realize first that firm values in Equations (7)-(8) are driven by changes in flow profits π_t and robot adoption costs c_t^R. Flow profits depend in turn on firm unit costs C_t, manufacturing demand Y_{Mt}, and the manufacturing

price P_{Mt}; see Equations (5) and (6). The predatory investment externality works through the price index P_{Mt}. When tabulating the effects on firm values in Table 5, I hold this externality fixed by calculating

$$\tilde{v}_t^T - v_t^B = v(\{c_\tau^{RT}, C_\tau^T, Y_{M\tau}^T, P_{M\tau}^B\}_{\tau=t}^\infty) - v(\{c_\tau^{RB}, C_\tau^B, Y_{M\tau}^B, P_{M\tau}^B\}_{\tau=t}^\infty), \tag{90}$$

where superscripts T and B denote the robot tax counterfactual and baseline equilibrium, respectively.

Table G.2 now incorporates the predatory investment externalities by calculating

$$v_t^T - v_t^R = v(\{c_\tau^{RT}, C_\tau^T, Y_{M\tau}^T, P_{M\tau}^T\}_{\tau=t}^\infty) - v(\{c_\tau^{RB}, C_\tau^B, Y_{M\tau}^B, P_{M\tau}^B\}_{\tau=t}^\infty) \tag{91}$$

Table G.2 shows a stark finding: For baseline values of model parameters, the predatory externalities are large enough to make total tax revenues exceed total profit losses from the robot taxes. Put differently, if tax revenues can be rebated to firms appropriately, a robot tax has the potential to increase firm values by internalizing the predatory externalities of robot adoption.

Table G.2: Robot Tax Incidence with Predatory Investment Externalities
(Discounted Present Values in Percent of GDP in 2019)

	Temporary Tax	Permanent Tax
Profits	-1.65	-7.90
Predatory Investment Externalities	2.48	2.67
Tax Revenues	2.39	9.41

I hold these predatory externalities on firm profits out of the baseline analysis to focus

on the key equity-efficiency trade-off for workers. That said, the analysis in this section suggests that studying predatory implications of recent automation technologies may be a fruitful avenue of future research.

Chapter 2

Individual and Market-Level Effects of Retraining Subsidies:

Evidence From a Natural Experiment

A version of this chapter was previously presented for the Danish Agency for Labour Market and Recruitment.

1 Introduction

Rising pressures from automation and globalization have put retraining policies in the spotlight. For example, in 2019, the Annual Meeting of the World Economic Forum led to calls for a "reskilling revolution" to secure a "future of jobs for all" (World Economic Forum, 2020).

Yet, despite the global reach of new technologies and import competition, governments around the world currently pursue widely differing strategies with respect to worker retraining. In 2010, the United States spent only 0.04 percent of GDP on training programs. In comparison, Denmark spent almost 20 times that amount.[1]

In this paper, I ask how retraining subsidies impact training and labor markets. First, do subsidies boost retraining activity? Second, do retraining courses help workers find jobs? To answer these questions, I leverage a large reform of retraining subsidies for unemployed workers in Denmark. Implemented in January 2011, the reform unexpectedly eliminated all subsidies for 40 percent of the course activity. The subsidy cuts took the form of a price cutoff, whereby course prices above the cutoff lost their entire subsidy.[2] The reform eliminated subsidies for more than 800 different courses, including truck and bus driver licenses, and certificates for heavy equipment and welding machines. The price cutoff was repealed in April 2012 by a newly elected social-democratic government.

I use this natural experiment to estimate the individual and market level impacts of

[1] Appendix Figure A.1 shows the public expenditures on adult vocational training across OECD countries. For comparison, public expenditures on adult vocational training in Denmark (0.7 percent of GDP) are similar to spending on college and university education (Hummels et al., 2012).

[2] Course prices were prefixed on government budgets, preventing course suppliers from lowering their prices below the cutoff.

retraining subsidies. First, studying the unemployed workers directly affected by the reform, I find that retraining decisions respond strongly to subsidies. Participation in courses above the cutoff falls almost one-to-one with the subsidy program, and I find very limited substitution toward courses below the price cutoff. Second, I find clear evidence that training helps workers find jobs. When the subsidies are cut, unemployed workers suffer depressed job finding rates, especially in occupations that are targeted by courses above the cutoff. Third, studying training markets, I find that retraining subsidies are effective in boosting total course activity. Once the subsidies are cut, firms and workers do not pay for the training courses out of own pocket. Finally, I study how the subsidy cuts impacted the labor market for truck and bus drivers, the two occupations most severely affected by the reform. I find that bus companies responded to the lower inflow of licensed bus drivers by retaining their incumbent drivers. However, I do not find evidence that bus or trucking companies offered higher starting wages to attract new drivers during the cutoff period.

This paper contributes to a literature that evaluates active training programs (Heckman et al., 1999; McCall et al., 2016). This literature has identified two key identification challenges. The first challenge is that workers self-select into training based on their future job prospects, making it difficult to infer the causal impact of training using observational data (Ashenfelter, 1978).[3] The second challenge is that training programs could have equilibrium impacts in training or labor markets, and that such equilibrium effects are not identified in experiments that rely on worker-level randomization.[4]

[3]For example, we may observe that many workers who obtain crane certificates subsequently transition into the construction sector. However, these workers may have obtained their certificate exactly because they had favorable job opportunities in the construction sector to begin with.

[4]For example, in 2019, Amazon committed $700 million to actively retrain a third of their US work-

The subsidy reform constitutes a natural experiment that allows me to overcome the two identification challenges. First, at the worker level, the availability of retraining subsidies changes discontinuously for workers who lose their job around the reform dates. Second, the large scale of the reform provides us with a unique opportunity to assess how retraining subsidies impact equilibrium outcomes in training and labor markets.[5]

As for equilibrium impacts, retraining subsidies could crowd out private training efforts that firms or workers might otherwise have undertaken, or create "rat races" in labor markets wherein trainees simply displace other qualified workers. Indeed, recent research have established the empirical relevance of rat race effects for other active labor market policies, including job search assistance (Crépon et al., 2013) and unemployment benefits (Lalive et al., 2015). However, it is not clear that experiences from job search assistance can be extrapolated to understand the labor market impacts of retraining subsidies. In fact, a goal of retraining programs might be to help workers transition from occupations with excess supply towards those with excess demand, thus mitigating the potential for rat race displacement effects.[6] McCall et al. (2016) review the literature and conclude that *"we have not found a single study of vocational training that takes equilibrium considerations seriously"*.[7] This paper provides the first causal evidence on the market-

force. A key counterfactual question is whether Amazon would have undertaken a similar investment had retraining subsidies in the United States been as generous as in Denmark.

[5]Prior to the reform, the subsidy program accounted for 35 percent of all new licenses for trucks and buses in Denmark.

[6]The stated goal of the retraining subsidy program in Denmark is to *"solve labor market restructuring and adaptation problems in accordance with demands on the labor market in the short and long run."* (Danish Ministry of Education, 2018).

[7]The full quote reads *"We have not found a single study of vocational training that takes equilibrium considerations seriously. The absence of evidence on the equilibrium effects of government-sponsored vocational training for adults represents a real opportunity for policy-relevant and methodologically challenging research. And, of course, as the old saying goes, the absence of evidence is not evidence of the absence, in this case of substantively important general equilibrium effects in the training context."*

level impacts of retraining subsidies. However, due to its short life from January 2011 to April 2012, the subsidy reform mainly allow us to learn about the short-run effects of retraining subsidies.

Finally, to be clear, this paper will not answer the normative question of whether governments *should* subsidize retraining.[8] Rather, I will focus on positive questions whose answers are key inputs into any normative analysis of the role for retraining subsidies.

The remainder of the paper is organized as follows. Section 2 lays out the institutional setting of the retraining subsidy reform. Section 3 describes my sources of data. Sections 4 and 5 use the reform to estimate the individual and market level effects of retraining subsidies. Section 6 concludes.

2 Institutional Setting

2.1 Retraining for Unemployed Workers in Denmark

In Denmark, unemployed workers are eligible for six weeks of vocational training free of charge.[9] Workers choose from a catalogue of about 2,500 courses and continue to receive unemployment benefits while training. Each training course has a list price that providers charge the government. For example, a truck driver license is a relatively expensive course listed at 5,825 Kroners ($1,080) per week for 30 work days. A course in elderly care for dementia patients is less expensive listed at 2,185 Kroners ($405) per week.

[8]Motives for retraining subsidies include market failures (credit constraints, incomplete insurance contracts, knowledge spillovers, frictions in labor markets, etc.) or targeted redistribution (for example, giving workers that have been displaced by automation or globalization a second chance in the labor market).

[9]The program (called "6-ugers selvvalgt uddannelse" in Danish) is eligible for unemployed workers who are insured through an unemployment fund in Denmark ("dagpengemodtagere").

The course prices are set annually on the government budget.

2.2 The Price Cutoff Reform

On January 1, 2011, a strict price cutoff of 3,500 Danish Kroners ($650) per week was put on training courses eligible for subsidies. For example, crane certificates were just 48 Kroners ($9) too expensive and thus lost its *entire* subsidy. Unemployed workers could still take the crane certificate, but would then need to pay the full list price out of own pocket. Other popular courses above the cutoff included truck and bus driver licenses, welding certificates, as well as several certificates for heavy equipment.

Table 1: Most Popular Courses in 2010

	Price per week (DKK)	Share in training (%)
Above Price Cutoff		36.8
Truck driver license	5,747	13.9
Bus driver license	5,354	7.3
Forklift certificate	3,748	4.7
Welding certificate	5,183	3.9
Crane certificate	3,548	2.7
Below Price Cutoff		63.2
Spreadsheet calculations	2,448	27.9
Taxi driver license	3,452	4.6
Presentation skills	2,780	3.5
Security guard training	2,478	2.4
Personal development	3,170	2.2

Notes: This table shows the most popular training courses in 2010, before a price cutoff of 3,500 DKK per week was imposed. Column (1) shows the course list prices, and Columns (2) reports the share of the course category in total training activity. The course categories reflect an aggregation of course codes into 125 categories, each of which prepares for work in a set of target occupations; see Appendix B.2 for details.

Table 1 shows the most popular courses before the cutoff was imposed. In total, the five most popular course categories constituted almost 90 percent of the course activity above the price cutoff.

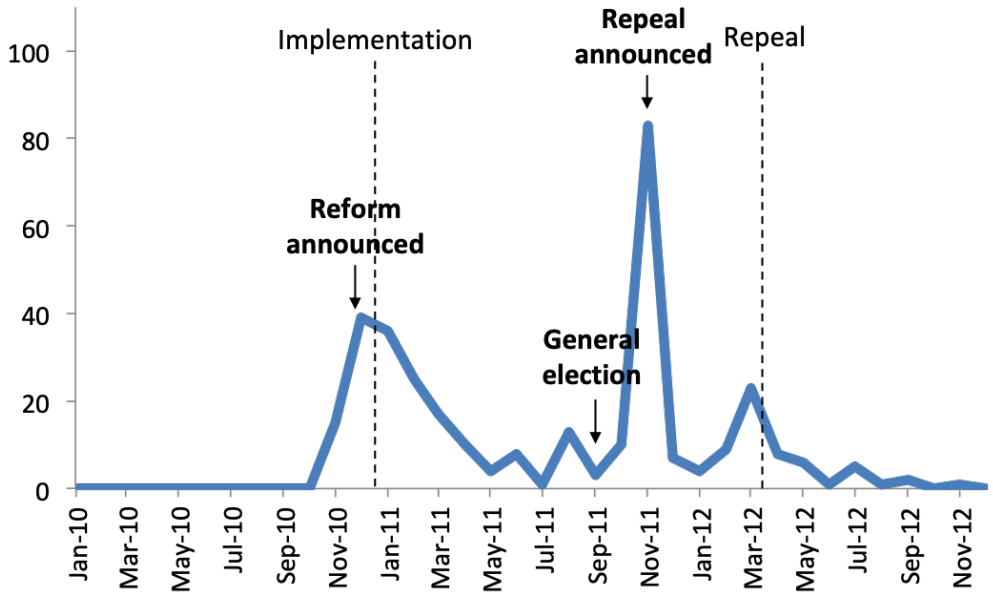
2.2.1 An Unanticipated and Quasi-Exogenous Reform

Figure 1 shows the monthly media coverage of the price cutoff reform. The reform was not mentioned once before the end of November 2010, when the then conservative government, in the dying minutes of an overnight negotiation, included it in a larger fiscal austerity package (Dagbladet Arbejderen, 2010).[10] A price cutoff was deliberately chosen so that the government would not have to evaluate the demands for each individual training course (Nordjyske Stiftidende, 2010). The specific cutoff of 3,500 Kroners was set to close a final gap of 1 billion Kroner in the government budget (Altinget, 2010).

The reform was originally planed to be permanent, but was highly unpopular among the labor movement in Denmark (Avisen.dk, 2011). Hence, shortly after coming into political power, a social-democratic government announced in November 2011 that a repeal of the price cutoff would come into effect on April 1, 2012.

[10]The price cutoff was included in the fiscal recovery package of 2010, "Genopretningspakken".

Figure 1: Newspaper Articles Mentioning the Price Cutoff Reform



Notes: This figure shows the number of newspaper articles containing the keywords: "prisloft selvvalgt uddannelse". Source: Infomedia.

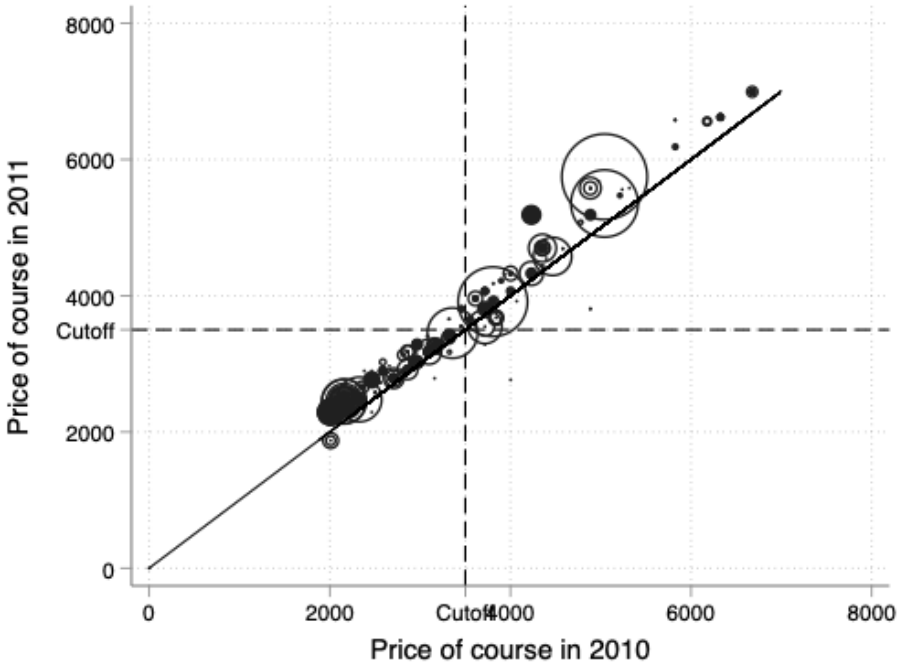
2.2.2 Government Budget Cycles Prevented Course Price Responses

The price cutoff reform created an incentive for course providers to lower their prices below the cutoff to requalify for the subsidies. In practice, however, government budget cycles prevented such price responses. Course prices are set around October in advance of each year. Thus, when the cutoff was first announced in November 2010, the course prices for 2011 had already been set.[11] Furthermore, when it came to negotiating the course prices for 2012, the political powers had changed, and course providers could plausibly expect the subsidies to be reinstated soon. As Figure 2 shows, we can almost perfectly

[11] A newspaper interview with a manager of one of the training facilities confirms this story: *"The schools are not allowed to lower the price or have the unemployed workers pay the price difference. "Trust me, we have looked into every possibility. We are on the government budget, and we are obliged by law to charge the set course price.""* (Politiken, 2011).

predict whether a course lost its subsidy in 2011 based its price in 2010.[12] Appendix Figure A.4 shows that course providers did not respond by creating new courses below the cutoff either.

Figure 2: Course Prices in 2011 versus 2010



Notes: This figure compares the prices for retraining courses in 2010 and 2011. The solid line represents the 45-degree line.

3 Data

I merge three administrative registers at Statistics Denmark: (*i*) a register of individual participations in vocational training courses, (*ii*) a database of government transfers to individuals, including retraining subsidy receipts and eligibility statuses, and (*iii*) a

[12] A mechanical prediction has a R-squared of 0.9974. The OLS prediction is

$$\mathbf{1}[P_{c2011} \geq 3500] = \underset{(0.0009)}{0.000} + \underset{(0.0015)}{0.9984} \times \mathbf{1}[P_{c2010} \geq 3500], \quad R^2 = 0.9974, \quad \bar{R}^2_{45} = 0.9974,$$

where $\bar{R}^2_{45}$ is the R-squared of the 45-degree line. Appendix Figure A.3 shows the corresponding figure for 2012.

matched employer-employee dataset. The resulting dataset is unusually rich, tracking all Danish workers across detailed training courses and in their employment relationships with precise information about occupation, hours, and wage rates on a monthly frequency.

Training data The Course Participant Register (VEUV) contains detailed information about individual participations in vocational training course, including 5-digit course codes, course supplier (private and public), training facility, and course dates. The database dates back to the 1970s. The register covers courses that are eligible for government subsidies, and then records all attendees of these courses regardless of their funding source.[13] In 2010, about 422,000 Danes (out of a labor force of 2.7 million) participated in one of the registered courses.

For several certificates, including those for heavy equipment and welding, the register contains the universe of formal training in Denmark. For driver licenses (taxa, bus, and truck), many – but not all – driving schools have registered to be eligible for government subsidies.Furthermore, for many of the covered courses, there is limited scope for informal training, either because it is unlawful (as with driver licenses or crane certificates), or because firm contracts require workers to have the formal certificates (e.g. metal suppliers that contract on certified welding standards).

I complement the Course Participant Register with information about course prices and course contents from government budgets ("Takstkataloger") and the Statistics Database of the Ministry of Education. In Appendix Section B.2, I group the training courses into

[13]The Course Participant Register covers vocational training, basic training, and post-secondary degrees ("diplomuddannelser"). I focus on vocational training courses (`ORGTILH` codes between 3000 and 4000).

125 categories, and link each of these course categories to its target occupations.

The Course Participation Register contains a unique person identifier for the training participant, allowing me to merge it with other administrative registers at Statistics Denmark.

Data on government transfers The DREAM (Danish Register for the Labor Market) Register constitutes a comprehensive account of government transfer payments to individuals on a weekly basis. I use this register to identify unemployment spells and to track the subsidy statuses of trainees. For example, since August 2009, the register contains an indicator for whether an individual participates in government-subsidized retraining for unemployed workers (the subsidy program affected by the price cutoff reform) in a given week. Appendix B.1 describes the measurement of unemployment spells.

Matched employer-employee data The Employer-Employee Register (BFL) contains the universe of job spells from January 2008 to December 2015. As informality in Denmark is very close to zero (Bentsen, 2018), the database covers virtually all paid work in the economy. The register contains worker identifiers, firm identifiers, salaries, hours, occupational codes, industry codes, as well as start and end dates for all job spells on a monthly frequency.[14] Finally, I pull socioeconomic data, such as age, gender, and education, from the Population and Education Registers (BEF and UDDA).

[14]The occupational data are drawn from the Wage Statistic (in Danish, "Lønstatistikken"), which covers all public sector jobs and all private sector firms with more than 10 employees. I use only the original recorded data of highest quality (`DISCO_MATCHPRIO_KODE` of 10, 20 or 30) that has been reported on the job spell within six months of the period of consideration.

3.1 Stylized Facts on Retraining in Denmark

In this section, I present three facts about retraining in Denmark that will guide the empirical analysis in Section 4. The first two facts document the selection of workers into training courses. The third fact concerns the timing of retraining during unemployment spells.

Fact 1. Low-Skilled Men Use Subsidies Above the Price Cutoff

Table 2 shows the training propensities of unemployed workers by their skill and gender. The table shows that it is primarily low-skilled men who use the subsidies for courses above the price cutoff. For workers who lost their jobs prior to the reform, 10.2 percent of low-skilled men used subsidies above the price cutoff. In contrast, almost no high-skilled women trained above the price cutoff. In Section 4, I use this heterogeneity in the take-up of subsidies to conduct differences-in-differences analyses of the subsidy cuts.

Table 2: Subsidized Training Rates of Unemployed Workers (Prior to Reform)

		Training Rates	
	Share of Spells	Above Cutoff	Below Cutoff
Low Skill × Man	43.3	10.2	9.0
Low Skill × Woman	27.4	1.5	12.3
High Skill × Man	12.1	1.3	1.9
High Skill × Woman	17.2	0.0	1.8

Notes: This table show the share of retraining subsidy recipients ("training rates") by skill and gender of the unemployed workers. Shares and rates are measured in percent. High skill represent long post-secondary education. "Prior to Reform" refers to unemployment spells that start from September 2009 (start of sample) to July 2010 (half-year before the reform is implemented).

Fact 2. Workers Sort Into Training Courses That Complement Their Work Histories

Table 3 shows the rates of training for the five most popular course categories above the price cutoff. The table is split according to whether the worker previously has worked in the occupation that the training course is targeting, a status I call an *occupational insider*.

The table delivers two insights. First, many workers train for occupations that they previously have held jobs in. For example, among previous truck drivers (first row of Panel (b)), 18 percent participate in further truck driver training. Second, occupational insiders mainly take advanced certificates (e.g. long combination vehicle (LCV) certification for truck drivers), while occupational outsiders take licenses (e.g. for bus or truck) or introductory non-certificate courses (e.g. for welding). In Section 4, I estimate the returns to vocational training for occupational insiders and outsiders separately. I also leverage the sorting of occupational insiders into training to test whether workers select into unemployment based on the availability of training subsidies.

Occupational insiders constitute on average 20 percent of the subsidized course activity. However, even among occupational outsiders, workers tend to reskill for occupations that are similar to jobs they have held in the past. For example, among low-skilled men who have not been bus drivers previously, past truck drivers are five times more likely to take a bus driver license. Appendix Table B.3 reports the relative course participation rates of occupational outsiders according to their previous jobs.

Table 3: Training Rates of Occupational Insiders and Outsiders (Prior to Reform)

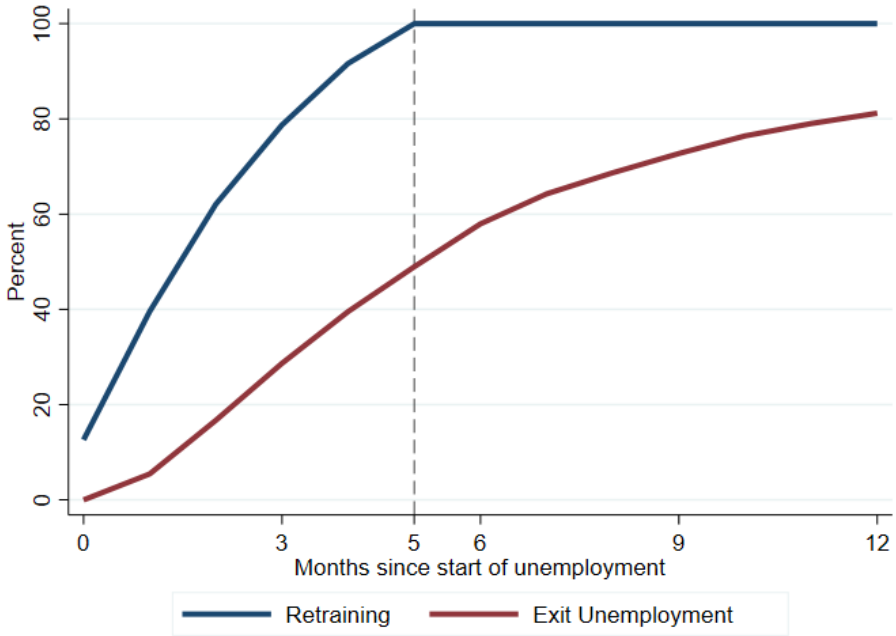
	Unemployed Workers		Trainees		
	Share of Spells	Training Rate	Licenses	Certificates	Other
			(a) *Bus Driver Training*		
Insiders	0.8	9.3	7.1	87.5	5.4
Outsiders	99.2	0.5	81.4	14.9	3.6
			(b) *Truck Driver Training*		
Insiders	3.6	13.9	1.8	97.9	0.3
Outsiders	96.4	1.1	52.7	45.7	1.6
			(c) *Welder Training*		
Insiders	4.3	9.7	0.0	87.7	12.3
Outsiders	95.7	0.4	0.0	73.5	26.5
			(d) *Crane Operator Training*		
Insiders	0.1	11.7	0.0	90.9	9.1
Outsiders	99.9	0.7	0.0	91.7	8.3
			(e) *Forklift Operator Training*		
Insiders	6.8	2.2	0.0	100.0	0.0
Outsiders	93.2	2.2	0.0	99.9	0.1

Notes: This table reports the training rates for the five most popular course categories above the price cutoff. An occupational insider is defined as a worker who previously has held a job in the occupation that the training course is targeting (for example, a previous welder taking a welding certificate).

Fact 3. Workers Retrain Within Six Months of Becoming Unemployed

Figure 3 shows the share of subsidized trainees who have started their training (blue line) or exited unemployment (red line) by the number of months since they became unemployed. The figure shows that all workers start their retraining within six months of losing their job. This training window is mandated by law, and implies that the date a worker loses his job determines whether he is exposed to the subsidy reform.[15] In Section 4, I leverage this timing fact to estimate how retraining subsidies impact individual unemployed workers.

Figure 3: Cumulative Hazard Functions for Retraining Subsidy Recipients



Notes: This figure shows the share of workers who have started retraining or exited unemployment by the number of months since they became unemployed. The hazard rates are calculated among workers who received subsidies for retraining courses above the cutoff.

[15]Workers older than 25 years have to complete any subsidized training within nine months of unemployment (Ramboll, 2011). For workers under the age of 25, the subsidy eligibility window is six months.

4 Individual-Level Effects of Retraining Subsidies

In this section, I use the subsidy reform to estimate the causal impact of retraining subsidies on the unemployed workers that are eligible for subsidies. I first study how the reform impacted retraining activity in Section 4.1, and then turn to estimate its effects on labor market outcomes in Section 4.2.

4.1 Retraining Activity

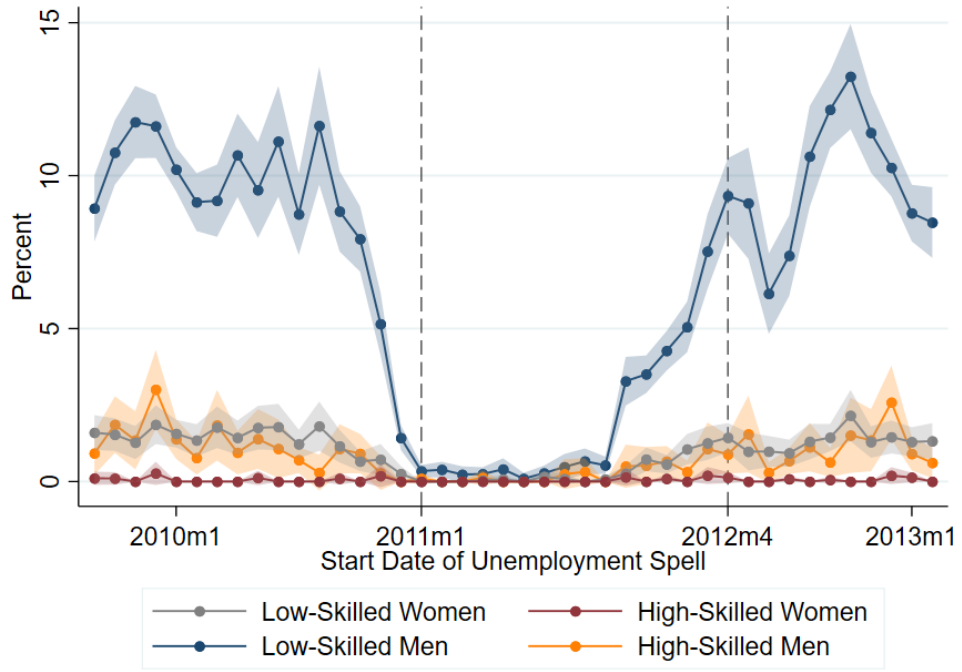
To identify how subsidies affect individual propensities to retrain, I adopt a differences-in-differences specification that leverages that the timing of unemployment (Fact 3), together with worker observables (Facts 1 and 2), predict a worker's exposure to the subsidy reform. In Section 4.1.1, I first document the first-stage impact of the price cutoff reform on the subsidized retraining activity above the cutoff. I then investigate whether the subsidy program crowds out other funding sources, private or public, for retraining in Section 4.1.2. In Section 4.1.3, I study how workers substitute between retraining courses in response to the altered subsidy structure caused by the reform. Put briefly, I find that retraining activity in eligible courses respond strongly to subsidies, and that workers show very limited ability to substitute to either other retraining courses or alternative funding sources.

4.1.1 Subsidized Retraining Above the Price Cutoff

In Figure 4, I plot the share of workers that receive subsidies for training courses above the cutoff by start date of their unemployment spell. As the figure shows, the date a worker

loses his job, together with his skill and gender, determines his use of retraining subsidies. For low-skilled men, the most exposed group (Fact 1 of Section 3.1), the subsidy rate dropped from around ten to zero percent when the cutoff was imposed. In comparison, the subsidy rate for high-skilled women remained close to zero throughout the analysis period. Appendix Table C.1 presents the corresponding regression estimates. These impacts represent the first-stage effects of the cutoff reform on the subsidized retraining activity above the price cutoff.

Figure 4: Subsidy Recipients for Courses Above the Cutoff by Start Date of Unemployment



Notes: This figure shows the share of workers that receive subsidies for retraining courses above the price cutoff by start date of their unemployment spell. Shaded areas represent 95 percent confidence bands.

In order to assign a causal interpretation to the estimates in Figure 4, an identifying assumption is that workers do not select into unemployment based on the availability of

training subsidies. Using the fact that workers select into retraining based on their work histories (Fact 2 of Section 3.1), we can test this "no selection" assumption against worker observables. In support of the "no selection" assumption, Appendix Table C.2 shows that the correlation between becoming unemployed during the reform period and being an occupational insider to a course above the price cutoff (the group with the highest training propensities, cf. Table 3) is a precisely estimated zero.

4.1.2 Substitution Across Funding Sources

Table 4 shows the impact of the subsidy reform on retraining activity across different funding sources. Column (1) reproduces the finding in Figure 4 that the subsidized training rates of low-skilled men fell about nine percentage points during the reform period. Column (2) shows how the reform impacted retraining above the cutoff funded by all other sources, including out-of-pocket expenses and other public programs. The column shows a small, but positive effect of the subsidy cuts on expenses from these alternative funding sources: Every dollar of federal subsidies crowds out 17 cents from other sources. In Column (3), I split out one of these alternative funding sources: a locally funded program called "job plans". Comparing Column (2) and (3), we see that the entire crowding-out effect comes through this local public program. Expenses from other sources, including private out-of-pocket expenses, were unaffected by the change in subsidies.

Table 4: The Impact of the Reform on Retraining Above the Cutoff by Funding Source

	Federal Subsidy	Other (incl. local)	Local Program
	(1)	(2)	(3)
		(a) *Pooled*	
Reform	-3.564	0.613	0.685
	(0.705)	(0.215)	(0.216)
		(b) *By Skill and Gender*	
Low Skill × Man × Reform	-9.295	1.757	1.922
	(0.689)	(0.479)	(0.493)
Low Skill × Woman × Reform	-1.371	0.089	0.118
	(0.229)	(0.084)	(0.026)
High Skill × Man × Reform	-0.806	0.141	0.192
	(0.203)	(0.133)	(0.077)
High Skill × Woman × Reform	-0.042	-0.023	-0.012
	(0.015)	(0.018)	(0.016)
Constant	4.269	1.101	0.190
Spells	200254	200254	200254

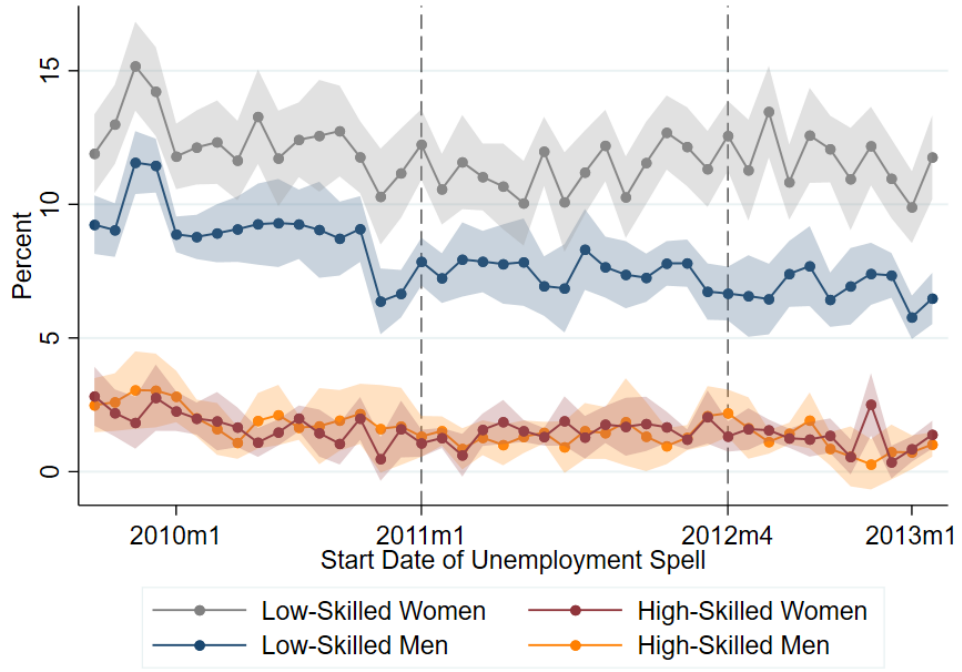
Notes: This table presents regressions of training above the cutoff by funding source on becoming unemployed in the reform period. Panel (a) shows the average effect, while Panel (b) splits the effects by gender and skill. Gender-skill-month fixed effects have been absorbed. Standard errors are clustered at the gender-skill and date (two-way) level.

Appendix Figure C.1 shows graphically the increase in local funding during the cutoff period. The importance of crowding out of substitute public programs mirrors the findings of Kline and Walters (2016) for early childhood education in the United States.

4.1.3 Substitution Toward Courses Below the Price Cutoff

Figure 5 shows how the subsidy reform impacted the subsidized retraining below the cutoff, i.e. in the courses whose subsidies were not cut by the reform. The training rates below the cutoff remained flat over the reform period, and Table 5 reports that the average effect was not significantly different from zero (Column (1) of Panel (a)).

Figure 5: Subsidy Recipients in Courses Below the Cutoff by Start Date of Unemployment



Notes: This figure shows the share of workers that receive subsidies for retraining courses below the price cutoff by start date of their unemployment spell. The series have been adjusted for (additive) seasonality. Shaded areas represent 95 percent confidence bands.

Why didn't more workers switch into courses below the price cutoff? One hypothesis is that the courses above versus below the cutoff simply do not represent close substitutes to one another. The most popular courses above the cutoff include large vehicle licenses

and heavy equipment certificates, while popular training activities below the cutoff include IT courses for office clerks (see Table 1).

To test the "no close substitutes" hypothesis, taxi driver licenses represent an interesting case study. Taxi driver licenses were not cut by the reform (their list price fell just 48 Kroners below the cutoff) and furthermore represent a close substitute for two of the major courses above cutoff, bus and truck driver licenses. In Column (2) of Table 5, I estimate the impact of the subsidy reform on subsidies for taxi driver licenses. Taxi driver licenses saw a significant increase during the cutoff period, reflecting a substantial amount of in-substitution from low-skilled men.

These estimates support the "no close substitutes" hypothesis: If more courses were like taxi driver licenses, we would have seen a larger substitution toward courses below the price cutoff.[16] An important take-away here is that workers do not opt into retraining courses that fall far from their past work experiences, even if the courses are heavily subsidized.[17]

[16]The popular press corroborates the "no close substitutes" story. From an article in the most popular Danish newspaper: *"Inside a classroom, 16 workers are taking a course in accounting. The participants are mostly unemployed women with office clerk backgrounds, and they have been able to choose a course that costs less than 3,500 Kroners per week. Accounting and IT courses are not affected by the price cutoff. However, these courses do not cater to the group of workers that take heavy equipment and transportation courses.* ***"I don't know what to use a course in accounting for. I have driven truck for 22 years"****, says Søren Jensen, driver for Statoil."* (Politiken, 2011).

[17]The "no close substitutes" story is consistent with Fact 2 of Section 3.1 that workers only use training courses that relate closely to their work histories.

Table 5: The Impact of the Reform on Subsidized Retraining Below the Cutoff

	Below Cutoff	Taxi license
	(1)	(2)
	(a) *Pooled*	
Reform	-0.278	0.091
	(0.150)	(0.032)
	(b) *By Skill and Gender*	
Low Skill × Man × Reform	-0.347	0.254
	(0.325)	(0.074)
Low Skill × Woman × Reform	-0.400	0.027
	(0.288)	(0.027)
High Skill × Man × Reform	-0.199	0.003
	(0.069)	(0.022)
High Skill × Woman × Reform	-0.069	0.000
	(0.125)	(0.000)
Constant	7.050	0.188
Spells	200254	200254

Note: This table presents regressions of subsidized training below the cutoff on becoming unemployed during the reform period. Column (1) includes all courses below the price cutoff. Column (2) focuses on taxi driver licenses. Panel (a) reports the average effect, while Panel (b) splits the effects by gender and skill. Gender-skill-month fixed effects have been absorbed. Standard errors are clustered at the gender-skill and date (two-way) level.

4.2 Labor Market Outcomes

In this section, I study how the price cutoff reform impacted labor market outcomes of unemployed workers. To recall, courses that lost their subsidy from the reform included truck and bus driver licenses, welding certificates, as well as several certificates for heavy

equipment. I evaluate these courses on their effects on employment and earnings, and on their ability to facilitate transitions into their target occupations. A key take-away from Section 4.1 is that the reform mainly affected the retraining activity low-skilled men. For this reason, I focus the labor market analysis on this sample of compliers. In summary, I find that training helps workers find jobs, and that these employment gains mainly accrue to workers that take licenses to bus and truck driving, and to former welders and truckers that participate in upskilling.

4.2.1 Employment and Earnings

To estimate the causal effects of retraining on labor market outcomes, I compare workers who lost their jobs under the reform period, when training subsidies were cut, to similar workers who lost their jobs either before or after the cutoff period. Following on this idea, Table 6 presents two-stage least squares (2SLS) estimates of the employment effects of participating in a retraining course above the price cutoff. I measure employment outcomes one year after a worker starts his unemployment spell.

Table 6 shows that retraining helps workers find jobs. Participating in a retraining course above the cutoff raises a worker's average hours of employment by more than 20 percent (Column (1)). This sizable employment effect reflects a higher job finding rate (Column (2)), as work hours conditional on employment are unaffected by the reform (Column (3)).

Where do the retrained workers find jobs? In Columns (4) to (6), I zoom in on the five main occupations targeted by courses above the cutoff: truck drivers, bus drivers,

welders, crane operators, and forklift operators.[18] The columns show that the increase in employment reflects extra jobs in these five target occupations, whose employment rates more than double upon retraining.

Table 6: Employment Effects of Retraining Courses Above the Price Cutoff (2SLS)

	All Occupations			Target Occupations		
	Monthly Hours	Employment Rate	Hours of Employed	Monthly Hours	Employment Rate	Hours of Employed
	(1)	(2)	(3)	(4)	(5)	(6)
Retrain Above Cutoff	21.8	17.8	-4.4	16.2	10.2	10.2
	(8.4)	(5.2)	(6.4)	(4.3)	(2.9)	(13.8)
Mean of Outcome	93.70	70.46	133.0	9.724	6.940	140.1
SD of Outcome	73.68	45.62	49.81	38.11	25.41	51.57
F stat	741.8	741.8	526.4	741.8	741.8	72.05
Observations	72482	72482	51124	72482	72482	4933

Notes: This table presents two-stage least squares estimates of the employment effects of retraining courses above the price cutoff. Columns (1)-(3) study employment in any occupation, while Columns (4)-(6) focus on employment in the five main occupations targeted by courses above the price cutoff. The regressions are estimated on the sample of low-skilled men, and employment outcomes are measured 12 months after start of unemployment. Retraining participation is instrumented with a dummy for starting unemployed during the reform period. Education-month fixed effects have been absorbed, where education represents six-digit categories and month is the start month of unemployment spell.

To assign a causal interpretation to the estimates presented in Table 6, an identifying assumption is that workers who lose their jobs under the reform period have similar potential employment outcomes as workers who become unemployed before or after the reform. In Appendix Table C.3, I perform a falsification test of this assumption, showing that workers' prior employment outcomes are uncorrelated with the reform treatment.[19]

[18] In Appendix B.2, I link each retraining course to its target occupations.

[19] The results of this placebo test are consistent with Appendix Table C.2, showing that the shares of occupational insiders to courses above the price cutoff were stable around the reform period.

To put the causal estimates in Table 6 into perspective, in Appendix Table C.4, I estimate the corresponding OLS relationships between retraining participation and employment outcomes. Interestingly, the causal employment gains from retraining are larger than the OLS correlations.

In Table 7, I turn to estimate the impact of retraining on earnings and wage rates. Although the earnings estimates are noisier, they show that the earnings gains from retraining mirror the positive employment effects estimated in Table 6. The impact of retraining on hourly wage rates conditional on employment is not statistically different from zero.

Table 7: Earnings Effects of Retraining Courses Above the Price Cutoff (2SLS)

	All Occupations		Target Occupations	
	Monthly Earnings (1)	Hourly Wage (2)	Monthly Earnings (3)	Hourly Wage (4)
Retrain Above Cutoff	571.7	-2.6	481.8	11.3
	(329.0)	(5.5)	(151.6)	(17.2)
Mean of outcome	3367.9	38.92	338.0	39.26
SD of outcome	2898.2	43.50	1349.4	65.95
F stat (first stage)	741.8	526.4	741.8	72.05
Observations	72482	51124	72482	4933

Notes: This table presents two-stage least squares estimates of the labor market earnings effects of participating in retraining courses above the price cutoff. Columns (1)-(2) study employment in any occupation, while Columns (3)-(4) focus on employment in the five main occupations targeted by courses above the price cutoff. Retraining participation is instrumented with a dummy for starting unemployed during the reform period. Earnings and wage rates are measured in US dollars.

How do the findings of this section compare to estimates in the existing literature? In their handbook chapter, McCall et al. (2016) conclude that *"the US literature suggests that*

zero is sometimes, but not always, a good summary of the impact of training programs." Furthermore, in the Danish context, a series of studies have documented negative correlations between training participation and subsequent labor market outcomes (Jespersen et al., 2008; Ramboll, 2011; Rosholm and Skipper, 2009; Sørensen et al., 2014).

There are at least two ways to reconcile the positive impacts estimated in this section with existing studies. First, whereas the existing literature consists almost entirely of observational studies (typically, using matching estimators), I leverage the subsidy reform to pursue a quasi-experimental research design. Indeed, comparing Table 6 to Appendix Table C.4, shows that the 2SLS estimates of the employment gains to retraining are larger than the OLS estimates. Second, the price cutoff reform impacted only a subset of courses (those with list prices above the cutoff), and these courses may have been particularly effective. In support of this hypothesis, in Appendix Table C.5, I estimate that participation in training courses *below* the cutoff correlates with lower employment rates and earnings. Furthermore, these negative impacts below the cutoff balance out the positive effects above the cutoff, such that the mean effect of retraining across all course types averages out to zero.

To better understand the positive returns generated by the courses above the cutoff, I next turn to investigate the effects of these courses separately.

4.2.2 Course-Specific Effects of Retraining

Table 8 shows the impact of the cutoff reform on the five most popular courses above the cutoff. For each of these courses, I estimate how the reform impacted training activity and job finding rates in their target occupations. Columns (3)-(4) and (5)-(6) estimate

these impacts separately for the groups of occupational outsiders and insiders.[20]

The table shows that the positive employment effects of training estimated in Column (5) of Table 6 reflects two sources of gains. First, for bus and truck driving, I find strong employment effects of training for occupational outsiders. As Table 3 showed, outsiders of bus and truck driving use the subsidies to take licenses, and Table 8 confirms that these outsiders do not become truck or bus drivers once the subsidies are revoked.

The second source of employment gains comes from occupational insiders in welding and truck driving. Although the estimates are noisy, the effects are consistent with the view that continuous upskilling, such as advanced certificates for welders or advanced freight licenses for truckers, is necessary for workers to stay employable in their occupations.

To assign a causal interpretation to the estimates presented in Table 8, an identifying assumption is that workers who lose their job under the reform period have similar potential occupational employment outcomes as workers who become unemployed before or after the reform. In Appendix Table C.6, I perform a falsification test of this assumption, showing that workers' prior occupations are uncorrelated with the reform treatment.

[20] *Occupational insiders* are workers who previously have held a job in the occupation that the training course is targeting. Table 3 shows descriptive statistics on the training behaviors of occupational insiders and outsiders.

Table 8: Training and Target-Job Employment Rates by Course Type (in Percent)

	All		Outsiders		Insiders	
	Training (1)	Target Job (2)	Training (3)	Target Job (4)	Training (5)	Target Job (6)
(a) *Bus Driver Training*						
Reform	-0.791***	-0.096	-0.678***	-0.192***	-9.074***	0.826
	(0.073)	(0.073)	(0.070)	(0.053)	(1.419)	(2.428)
Constant	1.069	0.784	0.959	0.476	9.049	21.216
(b) *Truck Driver Training*						
Reform	-2.235***	-0.489***	-1.666***	-0.252***	-9.402***	-1.985
	(0.158)	(0.135)	(0.141)	(0.072)	(0.945)	(1.190)
Constant	3.484	1.859	2.602	0.578	14.069	17.061
(c) *Welder Training*						
Reform	-0.634***	-0.188	-0.546***	0.015	-2.145*	-2.399
	(0.133)	(0.186)	(0.091)	(0.069)	(0.979)	(1.911)
Constant	1.949	1.984	1.053	0.549	11.466	17.078
(d) *Crane Operator Training*						
Reform	-1.024***	-0.013	-0.998***	0.001	-12.683	-3.946
	(0.121)	(0.013)	(0.120)	(0.009)	(6.670)	(7.904)
Constant	1.743	0.040	1.720	0.015	12.008	12.291
(e) *Forklift Operator Training*						
Reform	-2.591***	-0.003	-2.669***	-0.068	-1.986***	0.003
	(0.121)	(0.148)	(0.133)	(0.124)	(0.308)	(0.873)
Constant	3.671	2.379	3.818	1.296	2.563	10.666

Notes: This table presents regressions of course-specific training and employment on becoming unemployed during the reform reform. Training and employment rates are measured in percent. Education-month fixed effects have been absorbed, where education represent six-digit categories and month is the start month of unemployment spell.

In Appendix Table C.7, I convert the reduced-form effects in Table 8 into 2SLS estimates for the course-specific employment effects of retraining. For outsiders of bus and truck driving, the causal employment gains to training are larger than the OLS correlations. The estimates imply that providing workers with a truck or bus driver license closes the entire gap in their likelihood of finding a job as bus or truck drivers relative to workers who have held these jobs in the past.[21]

Finally, it is worth noting that the subsidy reform impacted occupational insiders through two opposing forces. On the negative side, subsidies for the certificates needed for continuous upskilling were cut. On the positive side, the reform lowered the inflow of newly trained workers, thus benefitting insiders in these occupations through increased labor market tightness. In Section 5.2, I return to analyze whether the subsidy reform had such equilibrium impacts on labor markets.

5 Market-Level Effects of Retraining Subsidies

In this part of the paper, I study whether the subsidy reform had equilibrium impacts in training or labor markets. In Section 5.1, I start off at the training course level, and ask if the subsidies crowded out training activity of firms or other workers in the economy. In Section 5.2, I then turn to occupational labor markets, and ask how the lower inflow of retrained workers impacted employment and wages. Before proceeding with the analysis,

[21] Let β_{cs}^{2SLS} denote the estimate in Table C.7 for workers of status $s \in \{\text{Insider}, \text{Outsider}\}$ of course c. Let β_{0cs}^{RF} denote the constant terms in Table 8. Training for truck driving then closes

$$\hat{\beta}^{2SLS}_{\text{Truck,Outsider}} / (\hat{\beta}^{RF}_{0,\text{Truck,Insider}} - \hat{\beta}^{RF}_{0,\text{Truck,Outsider}}) = 15.106/(17.061 - 0.578) = 91.6\%$$

of the gap in likelihood of finding a job as bus driver relative to workers who have held this job in the past. For trucking, this number is 136 percent.

I want to emphasize that, due to its short-lived nature, the subsidy reform mainly teaches us about the short-run effects of retraining subsidies. In summary, I find that retraining subsidies are effective in boosting total training activity in the economy. Second, I find evidence that firms responded to the lower inflow of trained workers by keeping their incumbent employees during the cutoff period.

5.1 Training Markets

In this section, I analyze how the subsidy reform impacted the total training activity in different courses. To recap, in Section 4.1, I found that course activity of unemployed workers fell almost one-to-one with the subsidy cuts. However, in these same courses, many employed workers also attend. For example, a metal fabrication shop may have its welders take an advanced certification, or a trucking company may use the courses for continuous upskilling of its truckers. Firms are compensated by the government for the training activities, but these compensations were not subject to the price cutoff.[22]

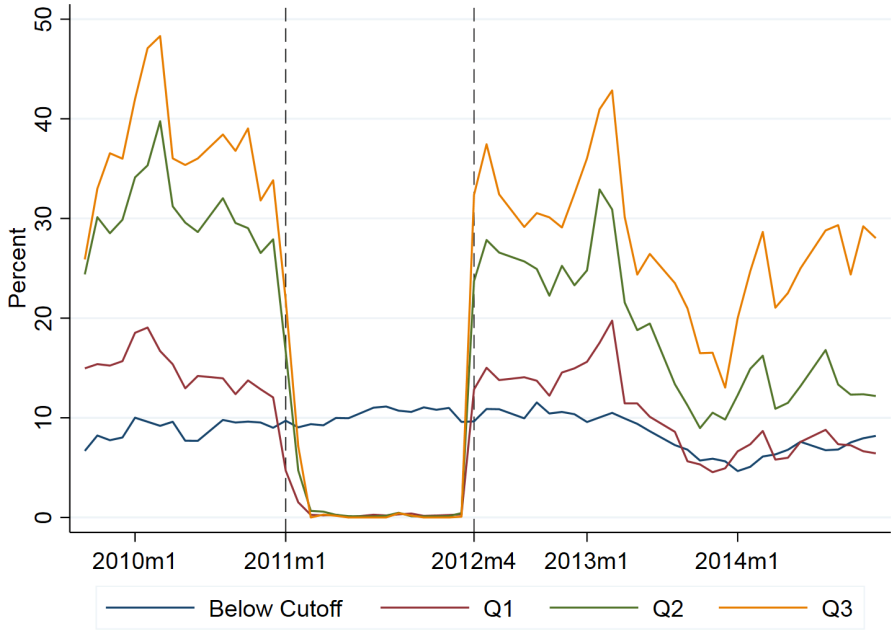
To study how firms and other workers responded when the training activity of unemployed workers was cut, I leverage two sources of variation. First, the reform only cut subsidies for courses above the price cutoff. Second, the importance of the subsidy program varied across different courses above the cutoff.

Following on this idea, in Figure 6, I split the courses above the cutoff according to the share of the subsidy program in total course activity in 2010. I then plot the evolution of the subsidy share over time across the course quantiles. The highest quantile (Q3)

[22]Humlum and Munch (2019) describe the system of vocational training for employed workers in Denmark.

predominantly constitute truck and bus driver licenses. For these most exposed courses, the subsidy share fell around 35 percentage points. In comparison, for the least exposed courses above the cutoff (Q1), the subsidy share fell 15 percentage points. Finally, it worth noting that the subsidy share of courses below the price cutoff remained stable throughout the reform period. This finding is in line with the take-away from Section 4.1.3 that unemployed workers displayed very limited substitution toward courses below the cutoff.

Figure 6: Share of Subsidy in Course Activity by Quantiles in 2010

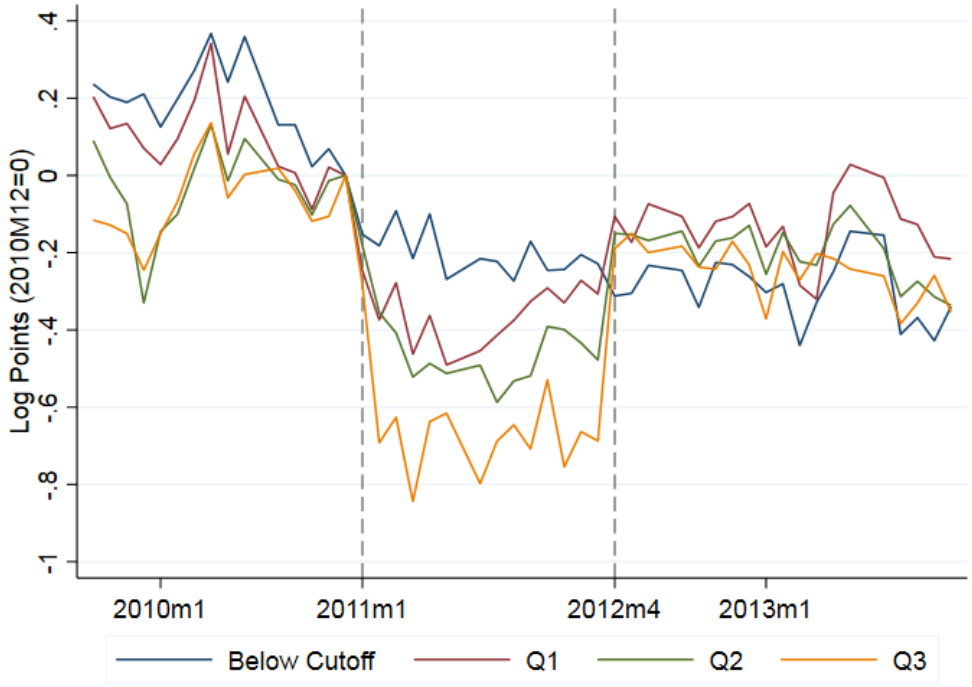


Notes: This figure shows the subsidy share in total course activity above and below the price cutoff. Training courses above the cutoff are divided into three quantile groups based on their subsidy shares in 2010.

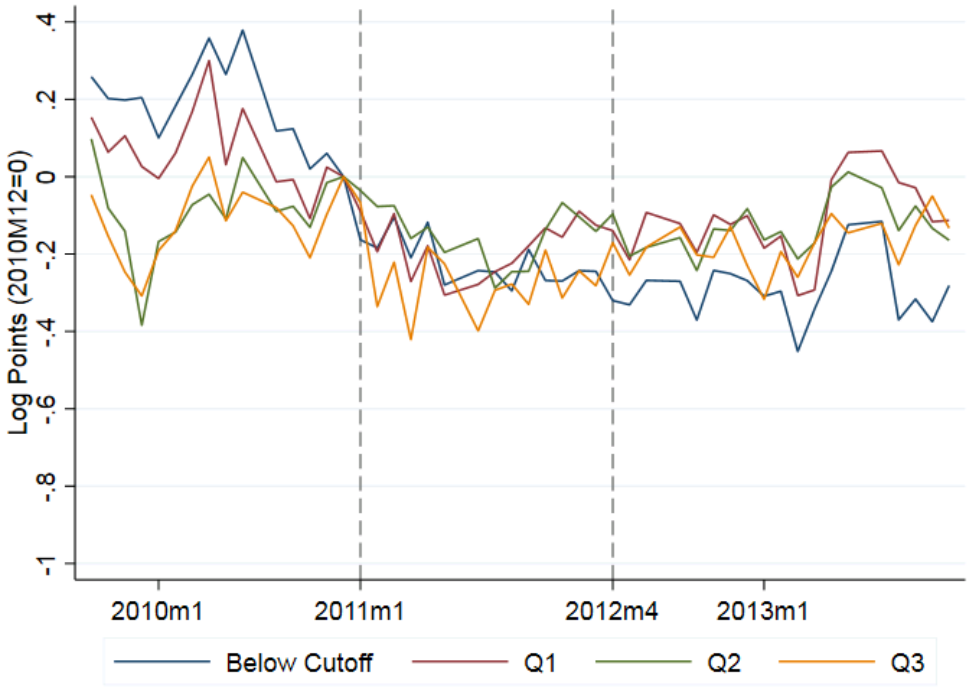
In Figure 7.(a), I plot the total training activity across the quantile groups. The figure shows that the training activity above the cutoff fell markedly during the cutoff period, and that the drops were steeper for courses with higher initial subsidy shares.

Figure 7: Course Activity by Quantiles of Subsidy Share

(a) Total Activity



(b) Non-Subsidized Activity



Notes: This figure shows the course activity above and below the price cutoff over time. Training courses above the cutoff are divided into three quantile groups based on their subsidy shares in 2010. Panel (a) plots the total course activity. Panel (b) focuses on the non-subsidized training activity.

However, by focusing on the non-subsidized training activity in Figure 7.(b), we see that the drop in training activity came from the subsidy program only. In fact, the non-subsidized activity was stable around the reform window, and did not respond differentially for courses that were more exposed to the reform through higher initial subsidy shares.[23]

Combining the takeaways from Panels (a) and (b), I conclude that retraining subsidies are effective in boosting total training activity in an economy. To sum up, I find that eligible workers respond strongly to subsidies, and I do not find that firms or other workers increase their training activities once subsidies are cut.

5.2 Labor Markets

In this final section, I ask how the lower inflow of retrained workers impacted occupational labor markets. For this purpose, I focus on the labor markets for bus and truck drivers for three reasons. First, bus and truck driving are the two occupations for which I find robust positive employment effects of retraining. Second, bus and truck driver licenses represent the two courses with the highest initial subsidy share.[24] Finally, bus and truck drivers have a close comparison group, taxi drivers, whose licenses were not cut by the subsidy reform.

In tracing out the responses of truck and bus companies to the lower inflow of licensed drivers, one of the most immediate margins is to retain incumbent drivers. In Section

[23]The non-subsidized course activity below the price cutoff exhibits a negative trend throughout the analysis window. This negative trend reflects a normalization from the financial crises, when many firms sent their employees on training courses to manage excess capacity (Danmarks Evalueringsinstitut, 2015).

[24]Appendix Figure D.1 shows that the share of recently subsidized bus and truck driver trainees represented more than 5 percent of all newly hired bus and truck drivers prior to the reform. This share dropped to zero during the cutoff period.

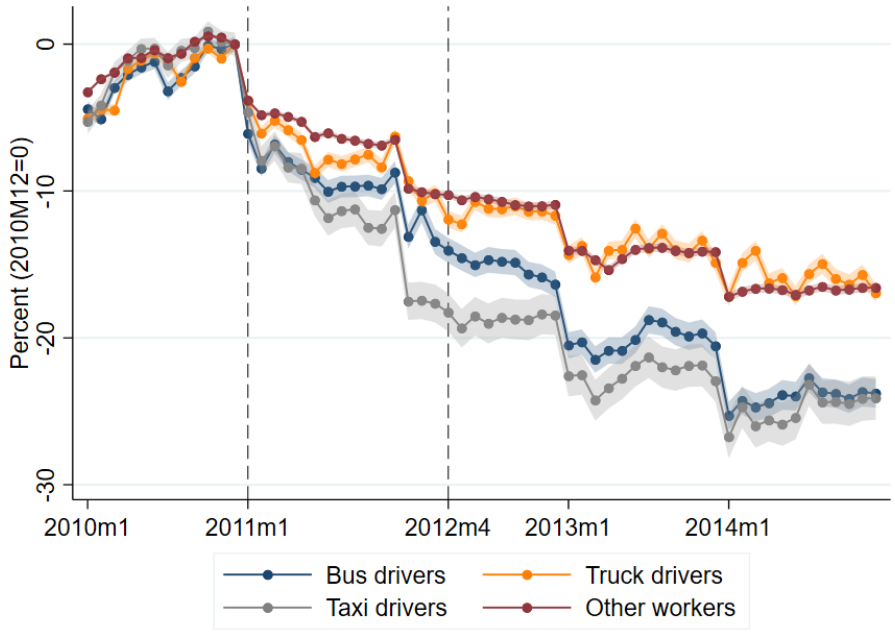
5.2.1, I study the employment outcomes of incumbent bus and truck drivers around the reform period. However, with the lower effective labor supply caused by the subsidy cuts, firms may have been forced to offer higher wages to attract new hires. In Section 5.2.2, I study the wage rates of newly hired drivers around the reform period.

Put briefly, I find that bus companies responded to the reform by retaining their incumbent drivers. However, I do not find evidence that new hires in truck or bus driving received higher starting wages during the reform period.

5.2.1 Employment of Incumbent Workers

In Figure 8, I follow workers who were full-time employed in different occupations in 2010. For each of these worker groups, I plot how their hours of employment evolve relative to their 2010 levels. Compared to the average worker, drivers generally have less steady ties to the labor market. While the employment of truck drivers track the average worker, bus and taxi drivers follow each other more closely. However, during the reform period, a gap of around five percentage points in employment opened up between bus and taxi drivers. The gap closed again after the training subsidies were reinstated. This temporary employment gap between bus and taxi drivers indicates that bus companies may have responded to the lower inflow of newly trained bus drivers by retaining their incumbent drivers.

Figure 8: Hours of Employment of Workers by Their Main Occupation in 2010



Notes: This figure shows hours of employment of workers relative to their levels in 2010. Workers are split according to their main occupation in 2010. I focus on workers who were at least 80 percent full-time employed in their main occupation in 2010. The time series have been adjusted for (additive) seasonality and then indexed to December 2010. Shaded areas represent 95 percent confidence bands.

5.2.2 Wages of New Hires

Table 9 estimates the impact of the cutoff reform on the wages of new hires in different occupations. The table shows that, relative to taxi drivers (the omitted occupation), new hires in truck or bus driving did not receive systematically higher wages during the cutoff period. Appendix Figure D.2 confirms this conclusion by plotting the monthly time series of starting wages in the different occupations.

Table 9: The Impact of the Price Cutoff on Wages of New Hires

	Log Wages (Percent)
	(1)
Bus driver × Reform	1.085
	(1.286)
Truck driver × Reform	0.140
	(1.161)
Other Occupations × Reform	0.440
	(1.018)
Observations	1763966

Notes: This table presents regression estimates of the impact of the price cutoff reform on the log hourly wage rates of new hires in different occupations. The effects are measured relative to taxi drivers, the omitted category. Hourly wage rates have been censored at 150 and 2,000 Danish Kroners. I focus of new hires in jobs with monthly hours equal to a half FTE or more. I estimate the regression on monthly data from January 2009 to December 2013, where the reform period extends from January 2011 through March 2012. Occupation-month fixed effects and a reform indicator have been absorbed.

These findings indicate that truck and bus companies did not respond to the lower inflow of trained drivers by offering higher wages to new hires. The fact that wages did not respond to retraining subsidies in the short run may have been due to nominal rigidities set in collective bargaining agreements.[25]

[25]In Denmark, the trade and labor unions that represent bus and truck drivers (3F-Transport and Dansk Industri) meet once every three years to set wage rates for the coming three years.

6 Conclusion

In this paper, I study a large reform of retraining subsidies in Denmark that delivers two headline insights. First, retraining subsidies are effective in boosting training activity in an economy. Second, retraining courses help workers transition between occupations. These findings indicate that retraining policies may help alleviate the disruption caused by automation or globalization on labor markets.

I find that training subsidies are particularly effective for workers that look to either transition into a licensed occupation or upskill in a former occupation through an advanced certification. In further work, I welcome further efforts to understand what types of courses that help workers find jobs. For example, a common characteristic of the courses affected by the reform is that they were relatively expensive certificates or licenses.

When left to own choice, I find that workers only opt in to courses that fall close to jobs they have held in the past. This suggests that workers may need more guidance if policymakers want to facilitate the more radical shifts of careers needed if a new technology or trading partner takes over a whole set of related tasks.

Finally, due to its short life from January 2011 to April 2012, the reform mainly teaches us about the short-run impacts of retraining subsidies. However, when policymakers discuss the need for a "reskilling revolution" (World Economic Forum, 2020), they are likely interested in the long-run equilibrium impacts of having retraining institutions put in place. In this paper, I take a first step toward understanding the equilibrium impacts

of retraining subsidies, but I encourage further work that aims to connect the short and long run consequences of retraining programs.

References

Altinget (2010). Støjberg skal spare en milliard.

Ashenfelter, O. (1978). Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics*, 47–57.

Avisen.dk (2011). Fagbevægelsen: Rød regering skuffer os.

Bentsen, K. H. (2018). *Aktiviteter uden for det formelle arbejdsmarked: sort arbejde, gør det selv-arbejde og deleøkonomi*. Syddansk Universitetsforlag.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora (2013). Do Labor Market Policies Have Displacement Effects? Evidence From a Clustered Randomized Experiment. *The quarterly journal of economics 128*(2), 531–580.

Dagbladet Arbejderen (2010). Farvel til 800 AMU-kurser.

Danish Ministry of Education (2018). Adult vocational training.

Danmarks Evalueringsinstitut (2015). Virksomheders brug og vurderinger af AMU i 2015.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999). The Economics and Econometrics of Active Labor Market Programs. In *Handbook of labor economics*, Volume 3, pp. 1865–2097. Elsevier.

Humlum, A. and J. R. Munch (2019). Globalization, Flexicurity and Adult Vocational Training in Denmark. *Making Globalization More Inclusive*, 51.

Hummels, D., J. R. Munch, L. Skipper, and C. Xiang (2012). Offshoring, Transition, and Training: Evidence From Danish Matched Worker-Firm Data. *American Economic Review 102*(3), 424–28.

Jespersen, S. T., J. R. Munch, and L. Skipper (2008). Costs and Benefits of Danish Active Labour Market Programmes. *Labour economics 15*(5), 859–884.

Kline, P. and C. R. Walters (2016). Evaluating Public Programs With Close Substitutes: The Case of Head Start. *The Quarterly Journal of Economics 131*(4), 1795–1848.

Lalive, R., C. Landais, and J. Zweimüller (2015). Market Externalities of Large Unemployment Insurance Extension Programs. *The American Economic Review*, 3564–3596.

McCall, B., J. Smith, and C. Wunsch (2016). Government-Sponsored Vocational Education for Adults. In *Handbook of the Economics of Education*, Volume 5, pp. 479–652. Elsevier.

Nordjyske Stiftidende (2010). Minister: Kurserne bliver billigere.

Politiken (2011). Prisloft sætter kraner og trucks i stå.

Ramboll (2011). Evaluering af 6 ugers selvvalgt uddannelse. *Report for the Danish Agency for Labour Market and Recruitment*.

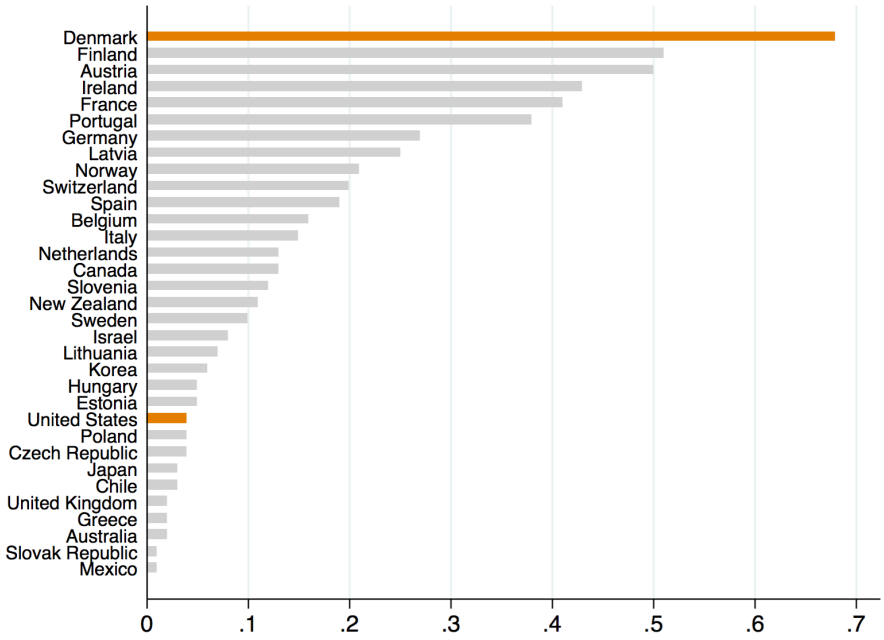
Rosholm, M. and L. Skipper (2009). Is Labour Market Training a Curse for the Unemployed? Evidence From a Social Experiment. *Journal of Applied Econometrics 24*(2), 338–365.

Sørensen, K. L., J. N. Arendt, and H. L. Andersen (2014). Effekter af uddannelsesaktivering for forsikrede ledige.

World Economic Forum (2020). Toward a Reskilling Revolution. www.weforum.org/projects/reskilling-revolution-platform. Accessed 2020-04-29.

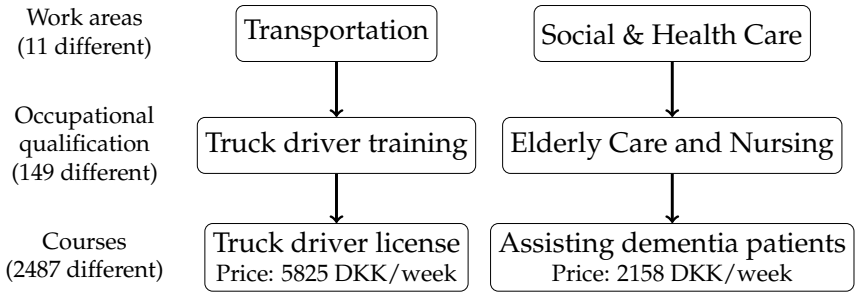
A Institutional Setting

Figure A.1: Public Expenditures on Training Programs in 2010 (Percent of GDP)



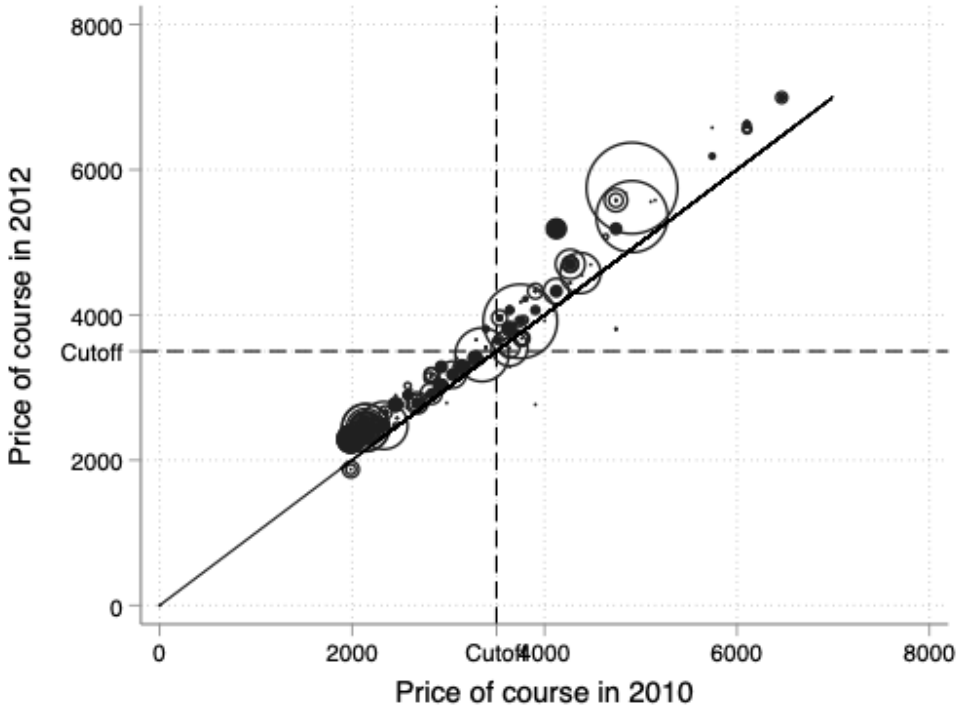
Notes: This figure shows the public expenditures on adult vocational training in percent of GDP across OECD countries. Source: OECD Employment Database.

Figure A.2: Structure of Retraining Courses



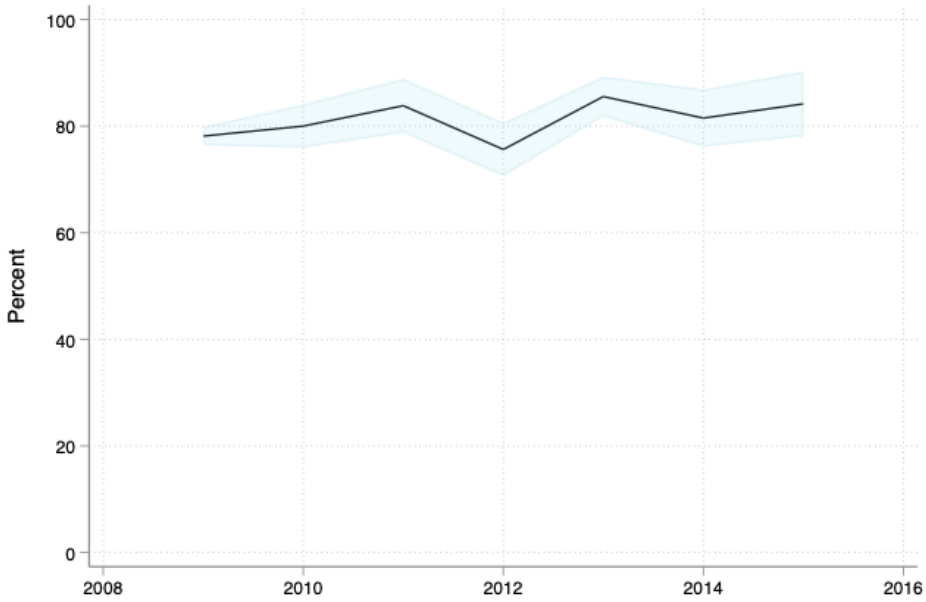
Notes: This figure illustrates the structure of retraining courses. The courses are organized into 149 occupational qualification descriptions ("fælles kompetence beskrivelser (FKB)") that lays out the occupational skills needed to carry out a specific job. The job qualification descriptions are in turn organized into eleven work areas ("efteruddannelsesudvalg").

Figure A.3: Course Prices in 2012 versus 2010



Notes: This figure compares the prices for retraining courses in 2010 and 2012.

Figure A.4: Share of New Courses That Have Prices Below Cutoff



Notes: This figure shows the share of new courses that have list prices below the cutoff of 3,500 DKK per week.

B Data

B.1 Measuring Unemployment Spells

In this section, I describe how the unemployment spells analyzed in Section 4 are measured. I define an unemployment spell as any consecutive months of unemployment benefit receipts. I focus on workers insured through an unemployment insurance fund ("dagpengeledige"), the group that is eligible for the retraining subsidy program.[26] The measurement steps 1 and 2 below follow directly from Ramboll (2011). Starting from the population of unemployment spells, I

1. *Unemployment spells starting 2009-2014*: Restrict to spells that represent new placements into the unemployment system.[27] This condition sorts out workers who frequently select in and out of the unemployment benefit system, and instead focus on workers who lost their job recently.
2. *Minimum one month of unemployment*: Exclude unemployment spells that are shorter than one month. This condition sorts out short unemployment spells that occur between job-to-job transitions.
3. *Data on worker characteristics*: Restrict to workers with information about age, gender, education, and occupational histories in the Employer-Employee Database (BFL) and the Employment Classification Module (AKM).
4. *Analysis workers*:

[26] Unemployed workers insured through an unemployment insurance fund correspond to DREAM codes 111-113, 121-126, 211-218, 231-232, and 299.

[27] I define a new placements as unemployment spells that start within one month of the recorded `INDT` (DREAM variable) date.

(a) Restrict to workers between the age of 25 and 55. The age limits ensure that labor market participation can be measured without interference from full-time studies or early retirement.

5. *Analysis Period*: For the diff-in-diff analysis in Section 4, focus on unemployment spells that start in either
   (a) The pre-period from September 2009 (start of sample) to July 2010 (half-year prior to reform implementation)
   (b) The reform period from January 2011 (reform implementation) to October 2011 (half-year prior to reform repeal)
   (c) The post-period from April 2012 (reform repeal) to February 2013 (end of sample)

Table B.1 shows the sample reduction table, where row "5. Analysis period" represents the 200,254 observations used in the regression analysis of Section 4.1. Of these unemployment spells, 72,482 are low-skilled men (the sample of workers studied in Section 4.2).

Table B.1: Measuring Unemployment Spells

Step	Spells at End of Step
1. Unemployment spells starting 2009-2014	816345
2. Minimum one month of unemployment	708634
3. Data on worker characteristics	565972
4. Analysis workers	429564
5. Analysis period	200254

B.2 Linking Training Courses to Target Occupations

To link training courses to their target occupations, I first group the 3912 different course codes offered between 2008 and 2015 into 125 course categories, each of which prepares for work in a set of target occupations. The course categories are based on the work area ("efteruddannelsesudvalg"), family group ("familiegruppe"), and price (above or below 3,500 DKK per week) of the training course. For a few course family groups, I disaggregate the course categories even further. I then link each of the 125 course categories to its target occupations (four-digit ISCO08 codes).

Table B.2 provides two examples of this mapping from training courses via course categories to target occupations. The full taxonomy of course-occupation links can be downloaded at `www.github.com/humlum/retraining`.

Table B.2: Linking Training Courses to Target Occupations: Two Examples

Training Course	Course Category	Target Occupations (ISCO-08)
Bus driver license	Bus driver training	Bus driver (8331)
TIG welding of steel pipes	Welder training	Welders and Flame Cutters (7212)
		Sheet Metal Workers (7213)
		Structural Metal Preparers (7214)
		Blacksmiths (7221)
		Toolmakers (7222)

B.3 Stylized Facts on Retraining in Denmark

Table B.3: Course Participation Rates of Occupational Outsiders (Top-3 Prior Occupations)

Course	Prior Occupation	Relative Training Rate $\frac{P(C\|O)}{P(C)}$
Bus driver training	Truck driver	5.0
	Taxi driver	4.1
	Mail carrier	3.3
Truck driver training	Bus driver	3.2
	Heavy equipment operator	2.4
	Motor vehicle mechanics	2.1
Welding training	Plumbers and pipe fitters	6.6
	Food machine operators	4.8
	Heating and cooling mechanic	4.0
Crane operator training	Truck driver	6.3
	Assemblers n.e.c.	3.1
	Blacksmiths	2.9
Forklift operator training	Food machine operators	2.0
	Sheet-metal workers	1.8
	Truck drivers	1.8

Notes: This table presents relative training rates by the previous occupation of an unemployed workers. Relative training rate is defined as the course participation probability relative to the average training rate in that course. The training rates are calculated among low-skilled men who have not before held a job in the target occupation (an 'occupational outsider').

C Individual-Level Effects of Retraining Subsidies

C.1 Retraining Activity

C.1.1 Subsidized Retraining Above the Price Cutoff

Table C.1: The Impact of the Price Cutoff on Subsidized Training Above the Cutoff

	Reform - Pre	Reform - Post	Pooled
	(1)	(2)	(3)
	(a) *Pooled*		
Reform	-3.894	-3.407	-3.564
	(0.320)	(0.346)	(0.207)
	(b) *By Skill and Gender*		
Low Skill × Man × Reform	-9.336	-9.308	-9.295
	(0.589)	(0.959)	(0.689)
Low Skill × Woman × Reform	-1.436	-1.279	-1.371
	(0.246)	(0.217)	(0.229)
High Skill × Man × Reform	-0.897	-0.753	-0.806
	(0.208)	(0.211)	(0.203)
High Skill × Woman × Reform	-0.021	-0.061	-0.042
	(0.014)	(0.030)	(0.015)
Constant	4.588	3.766	4.269
Spells	132601	126304	200254

Note: This table presents regressions of training above the cutoff by funding source on becoming unemployed in the reform period. Panel (a) shows the average effect, while Panel (b) splits the effects by gender and skill. Gender-skill-month fixed effects have been absorbed. Standard errors are clustered at the gender-skill and date (two-way) level. The reform period ranges from January 2011 to October 2011 (half-year prior to reform repeal). The pre-period ranges from September 2009 (start of sample) to July 2010 (half-year prior to reform implementation). The post-period spans April, 2012 (reform repeal) to February, 2013 (end of sample).

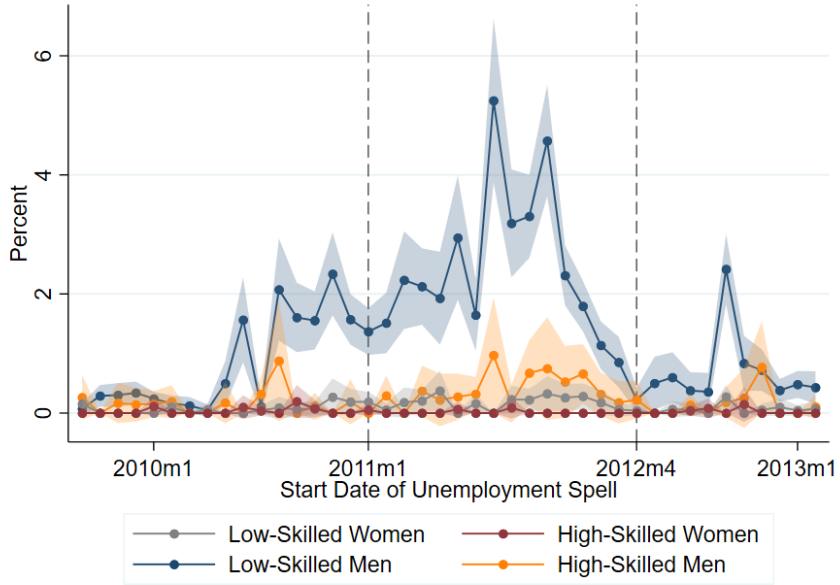
Table C.2: Falsification Test of Selection into Unemployment: Share of Occupational Insiders

	Any Insider	Bus Drivers	Truck Drivers	Welders	Crane Operators	Forklift Operators
Reform	-0.7	0.3	-0.5	0.3	-0.9	0.3
	(0.9)	(0.2)	(0.3)	(0.5)	(0.8)	(0.4)
Constant	10.9	0.9	3.5	4.5	0.1	6.5
Observations	200254	200254	200254	200254	200254	200254

Notes: This table presents regressions of being an occupational insider for one of the five most popular courses above the price cutoff on becoming unemployed during the cutoff period. The regression is estimated on the sample of low-skilled men, the group that is mainly exposed to the reform (Fact 1 of Section 3.1). Education-month fixed effects have been absorbed (education represents detailed six-digit categories) and standard errors are clustered at the education-month level.

C.1.2 Substitution Across Funding Sources

Figure C.1: Locally Funded Training Above the Cutoff



Notes: This figure shows the share of workers that participate in retraining courses above the price cutoff through the local program ("jobplans") by start month of their unemployment spell. Shaded areas represent 95 percent confidence bands.

C.2 Labor Market Outcomes

C.2.1 Employment and Earnings

Table C.3: Placebo Test: Prior Employment and Becoming Unemployed During the Reform

	All Occupations			Target Occupations		
	Monthly Hours	Employment Rate	Hours of Employed	Monthly Hours	Employment Rate	Hours of Employed
	(1)	(2)	(3)	(4)	(5)	(6)
Reform	-1.2	-0.6	-0.4	-0.2	0.1	-2.7
	(0.7)	(0.3)	(0.5)	(0.5)	(0.3)	(1.6)
Mean of outcome	93.70	70.46	133.0	9.724	6.940	140.1
SD of outcome	73.68	45.62	49.81	38.11	25.41	51.57
Observations	72482	72482	67849	72482	72482	7051

Notes: This table presents placebo regressions of a workers *prior* employment on an indicator for becoming the unemployed during the reform period. The regressions focus on low-skilled men, and employment outcomes are measured 12 months *prior* to start of unemployment. Education-month fixed effects have been absorbed, where education represents six-digit categories and month is the start month of unemployment spell. Furthermore, a dummy for the shift in occupational classification in 2010 (from ISCO88 to ISCO08) has been absorbed.

Table C.4: Employment Effects of Retraining Courses Above the Price Cutoff (OLS)

	All Occupations			Target Occupations		
	Hours	Employment Rate	Hours of Employed	Hours	Employment Rate	Hours of Employed
	(1)	(2)	(3)	(4)	(5)	(6)
Retrain Above Cutoff	5.2	3.3	1.0	10.6	7.2	4.2
	(1.0)	(0.6)	(0.8)	(0.5)	(0.4)	(2.2)
Mean of outcome	93.70	70.46	133.0	9.703	6.927	140.1
SD of outcome	73.68	45.62	49.81	38.11	25.41	51.57
Observations	72482	72482	51124	72482	72482	4933

Notes: This table presents OLS estimates of the employment effects of participating in retraining courses above the price cutoff. The regressions focus on low-skilled men, and employment outcomes are measured 12 months after start of unemployment. Gender-skill-date fixed effects have been absorbed, where education represents six-digit categories and month is the start month of unemployment spell.

Table C.5: Labor Market Effects of Retraining Courses (OLS)

	Hours	Employment Rate	Hours of Employed	Salary	Wage Rate
	(1)	(2)	(3)	(4)	(5)
			(a) *Pooled*		
Retrain	-0.3	-0.0	-0.3	-108.9***	-0.7
	(0.8)	(0.5)	(0.6)	(31.0)	(0.5)
			(b) *By Price Cutoff*		
Retrain Above Cutoff	-5.5***	-2.9***	-2.1*	-314.6***	-1.3
	(1.0)	(0.6)	(0.8)	(39.6)	(0.7)
Retrain Below Cutoff	6.1***	3.7***	1.3	200.3***	1.3
	(1.1)	(0.6)	(0.8)	(41.0)	(0.7)
Mean of outcome	93.70	70.46	133.0	3367.9	38.92
SD of outcome	73.68	45.62	49.81	2898.2	43.50
Observations	72482	72482	51124	72482	51124

Notes: This table presents OLS estimates of the employment effects of participating in retraining courses. Panel (a) include pools all retraining courses, while Panel (b) splits the courses above and below the price cutoff. The regressions focus on low-skilled men, and employment outcomes are measured 12 months after start of unemployment. Gender-skill-date fixed effects have been absorbed, where education represents six-digit categories and month is the start month of unemployment spell.

C.2.2 Course-Specific Effects of Retraining

Table C.6: Placebo Test: Prior Occupation and Becoming Unemployed During the Reform

	All	Outsiders	Insiders
	(1)	(2)	(3)
		(a) *Bus Driver Training*	
Reform	0.021	-0.005	0.575
	(0.119)	(0.005)	(4.744)
Constant	0.851	0.003	52.585
		(b) *Truck Driver Training*	
Reform	-0.104	0.017	0.548
	(0.176)	(0.017)	(1.802)
Constant	2.708	0.003	34.844
		(c) *Welder Training*	
Reform	0.243	-0.004	3.276
	(0.180)	(0.006)	(1.755)
Constant	2.643	0.004	30.036
		(d) *Crane Operator Training*	
Reform	0.004	0.000	12.058
	(0.028)	(.)	(17.960)
Constant	0.096	0.000	38.091
		(e) *Forklift Operator Training*	
Reform	-0.049	-0.011	2.889*
	(0.186)	(0.014)	(1.366)
Constant	3.441	0.017	28.263

Notes: This table presents placebo regressions of a worker's *prior* occupation on a dummy for becoming unemployed during the cutoff reform. The regressions focus on low-skilled men, and prior employment is measured 12 months *prior* to start of unemployment. Education-month fixed effects have been absorbed, where education represents six-digit categories and month is the start month of unemployment spell. Furthermore, a dummy for the shift in occupational classification in 2010 (from ISCO88 to ISCO08) has been absorbed.

Table C.7: The Impact of Training Course Participation on Employment Rate in Target Occupation

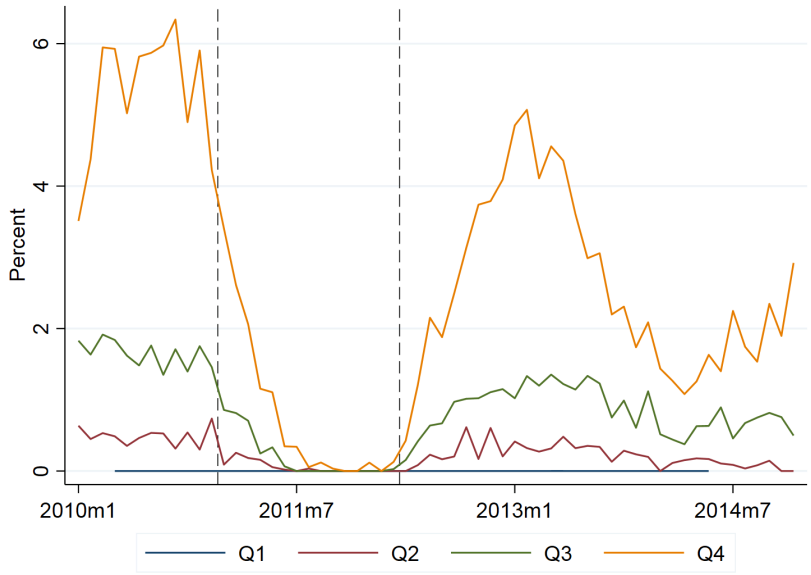
	Pooled		Outsiders		Insiders	
	2SLS	OLS	2SLS	OLS	2SLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
	(a) *Bus Driver Training*					
Course Participation	28.271	23.202	28.271	23.202	-9.106	18.550
	(8.431)	(0.266)	(8.431)	(0.266)	(36.935)	(7.216)
F stat (first stage)	69.06		69.06		22.33	
	(b) *Truck Driver Training*					
Course Participation	15.106	3.846	15.106	3.846	21.116	10.037
	(4.178)	(0.193)	(4.178)	(0.193)	(14.114)	(1.698)
F stat (first stage)	144.7		144.7		66.97	
	(c) *Welder Training*					
Course Participation	-2.806	5.930	-2.806	5.930	111.853	13.972
	(12.978)	(0.307)	(12.978)	(0.307)	(71.147)	(1.586)
F stat (first stage)	36.41		36.41		4.715	
	(d) *Crane Operator Training*					
Course Participation	-0.056	0.172	-0.056	0.172	31.111	3.597
	(1.136)	(0.039)	(1.136)	(0.039)	(88.208)	(15.616)
F stat (first stage)	82.13		82.13		1.858	
	(e) *Forklift Operator Training*					
Course Participation	2.545	2.184	2.545	2.184	-0.174	2.447
	(4.110)	(0.258)	(4.110)	(0.258)	(41.817)	(2.638)
F stat (first stage)	244.1		244.1		27.46	

Notes: This table presents two-stage least squares (2SLS) and ordinary least squares (OLS) estimates of the course-specific employment effects of retraining. Columns (1)-(2) show the pooled regressions, while Columns (3)-(4) and (5)-(6) split the regressions by whether the worker is an occupational insider for the training course in consideration. The regressions focus on low-skilled men, and employment outcomes are measured 12 months after start of unemployment. Gender-skill-date fixed effects have been absorbed, where education represents six-digit categories and month is the start month of unemployment spell.

D Market-Level Effects of Retraining Subsidies

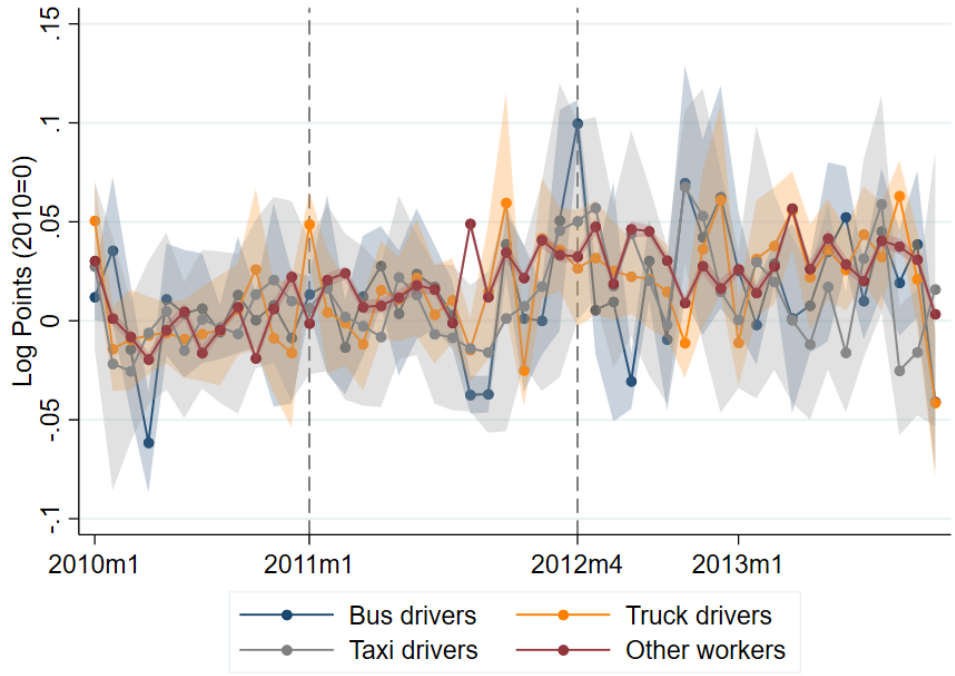
D.1 Labor Markets

Figure D.1: Share of Recently Subsidized Trainees in New Hires of Target Occupation by Quantiles in 2010 (Courses Above the Cutoff)



Notes: This figure shows the share of recently subsidized trainees in new hires of their target occupations. Occupations have been divided into quantile groups based on their subsidy shares in 2010. The fourth quantile (Q4), which represent bus and truck drivers, has the following interpretation: "Of all bus and truck drivers hired in 2010, five percent had obtained their license from the subsidy program within less than six months before starting their job".

Figure D.2: Hourly Wage Rate of New Hires



Notes: This figure shows logarithmic changes in wage rates of new hires by occupation. The time series have been adjusted for (additive) seasonality, and then indexed relative to the average of 2010. Hourly wage rates have been censored at 150 and 2,000 Danish Kroners. I focus of new hires in jobs with monthly hours of a half FTE (160.33/2) or more. Shaded areas represent 95 percent confidence bands.

Chapter 3

Artificial Intelligence and College Majors

A version of this chapter was previously presented at the University of Copenhagen.

1 Introduction

New technologies are a source of improved living standards, but may also pose risks of widening skill gaps and greater inequality. Today, Artificial Intelligence (AI) is one of the most salient technology arrivals, and its impact on labor markets is highly anticipated.

Artificial Intelligence, or Machine Learning, refers to algorithms that learn to complete tasks by identifying statistical patterns in data. AI is poised to be the next General Purpose Technology that will sweep through a wide range of industries (Brynjolfsson et al., 2018; Goldfarb et al., 2019). A distinctive feature of AI is that it may create winners and losers within the group of highly educated workers. While AI algorithms are being developed and implemented by specialized computer and data scientists, these same algorithms may be able to automate tasks done by other high-skilled workers, such as chemical engineers or optometrists (Webb, 2020).

In this paper, we document what types of college skills firms that use or produce AI employ. We view a college major as a set of skills that may be more or less relevant to work with AI. Using micro data from Denmark, we rank college majors according to their share of graduates that work in AI firms. Using these AI relevance rankings, we establish new facts about the college majors that work in AI firms, and we leverage an admission cutoff regression discontinuity (RD) design to estimate causal effects of pushing students toward more AI relevant degrees.

Our first key finding is that AI cuts through the group of Science, Technology, Engineering, and Mathematics (STEM) degrees. While computer science and mathematics

majors specialize in AI producer firms, we find that biology, chemistry, and other laboratory sciences concentrate in firms that only use AI. We find that AI producer relevance is correlated with higher earnings in the labor market and that these earnings premiums have increased in recent years. In contrast, we find that AI user majors do not earn higher wages, nor have experienced differential earnings growth recently.

Our second contribution is to provide causal evidence on an often discussed policy response to AI, that is to adjust the supply of different college majors (President's Council of Advisors on Science and Technology, 2012; Trajtenberg, 2018). In Denmark, admission to college majors is based on high school GPA cutoffs. These admission cutoffs naturally give rise to a regression discontinuity (RD) design, comparing outcomes of applicants around cutoffs for AI relevant majors. The RD analysis delivers two key insights. First, we find that it indeed is possible to push marginal students to obtain more AI relevant skills. Many students apply to majors of widely differing AI relevance, and the cutoffs create discontinuous jumps in admission rates to these programs. We find that students comply with these discontinuities, such that an applicant just above a cutoff has a discontinuously higher chance of graduating in that major. Second, using these discontinuities in college major attainment rates, we find that earnings premiums from random assignment to higher AI producer relevance are as large as those suggested by the correlation. In contrast, we find that the returns to higher AI user relevance is a precisely estimated zero. These estimated causal effects are informative about policies that aim to increase the AI relevance of college students by adjusting the admission capacity of different college majors.

We make a methodological contribution in this paper by proposing new measures of

the exposure of college majors to AI. A key advantage of our measures, which we micro-found in task-based theories of AI, is that they rely on revealed and objective information only. We correlate our revealed measures with characteristics of the college majors to shed light on the types of skills that are relevant for AI. First, using linked patent-occupation data from Webb (2020), we find that application-specific breakthroughs in AI are a key driver of the AI user relevance of different college majors. Second, we find that a core in programming or mathematics is a necessary ingredient for work with AI production. We do not find evidence that interpersonal skills alone qualify for work with AI, as social science and the humanities rank as the least relevant majors for AI firms. Our measurement approach does, however, require very detailed micro data on both AI technology by firms and the college majors held by workers. We leverage here the unusually rich data available in the Danish setting, where we can link register data on college diplomas of workers with a novel survey of AI production and use by firms.

This paper combines insights from two strands of literature. The closest related work is a recent set of papers that provide evidence on how AI relates to the demand for worker tasks (Alekseeva et al., 2019; Goldfarb et al., 2019; Webb, 2020). We make two contributions to this line of work. First, we use new micro data from Denmark to rank college majors according to their relevance for AI firms. Second, we adopt an admission RD design to estimate the causal returns to AI relevant skills. Admissions cutoffs have been used successfully in the literature on the returns to college field choice (Hastings et al., 2013; Kirkeboen et al., 2016). We complement these papers by linking college majors to a specific technology using a continuous measure that goes beyond traditional divisions of fields of study.

The remainder of the paper is structured as follows. Section 2 explains how we rank college majors according to AI relevance. Section 3 uses the AI relevance rankings to document facts about the college majors employed in AI firms. Section 4 conducts the admission regression discontinuity analysis. Section 5 concludes.

2 Measuring the AI Relevance of College Majors

In this section, we rank college majors according to their share of graduates that work in AI firms, a measure we call "AI relevance". To measure AI relevance, we link a novel survey of AI production and use by firms (described in Section 2.1) to register data on college diplomas of workers (described in Section 2.2). In Section 2.3, we relate the AI relevance measures to task-based theories of AI.

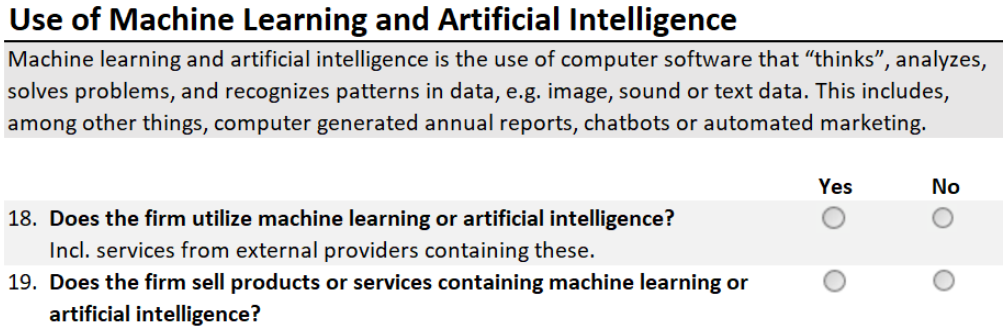
2.1 Firm Survey on Artificial Intelligence

Our micro data on AI comes from the representative VITA firm survey conducted by Statistics Denmark in 2017. The survey questions were prepared in collaboration with the Danish Business Authority as a supplement to Eurostat's technology survey. The survey sampled 3,890 firms from the population of private non-agricultural, non-financial firms with more than 10 employees. The response rate was 97 percent, and the surveyed firms covered 27 percent of total employment in Denmark.

Figure 1 shows the questionnaire on Artificial Intelligence. AI is defined as "computer software that "thinks", analyzes, solves problems, and recognizes patterns in data", and examples include computer-generated annual reports, chatbots, or automated marketing.

Firms were asked if they used AI in-house and whether they sold products or services containing AI technology.

Figure 1: Firm Survey on Artificial Intelligence



Use of Machine Learning and Artificial Intelligence

Machine learning and artificial intelligence is the use of computer software that "thinks", analyzes, solves problems, and recognizes patterns in data, e.g. image, sound or text data. This includes, among other things, computer generated annual reports, chatbots or automated marketing.

	Yes	No
18. **Does the firm utilize machine learning or artificial intelligence?** Incl. services from external providers containing these.	○	○
19. **Does the firm sell products or services containing machine learning or artificial intelligence?**	○	○

Among the surveyed firms, 2.2 percent answered 'yes' to selling AI-augmented products or services. AI producers employ 5.9 percent of workers, and 13.7 percent of college graduates. While most AI producers also use AI (70 percent), only a third of AI users sell AI-augmented products. For this reason, we will distinguish between firms that produce AI (henceforth, "AI producers") and those that use, but do not produce, AI (henceforth, "AI users"). Among the surveyed firms, 3.0 percent use, but do not produce AI. The AI users employ 10.3 percent of workers, and 18.9 percent of college graduates. Warzynski (2020) provides additional descriptives on the sample of AI using firms, studying how AI adoption correlates with firm industry, productivity, and workforce composition. In this paper, we use the AI survey to rank college majors according to their relevance for AI firms, which we turn to next.

2.2 Linking the AI Survey to College Diplomas

To calculate AI relevance, we first merge the AI survey to the employer-employee link file of the Integrated Database for Labor Market Research (IDAN). IDAN links the population

of workers to the universe of firms in November of each year, allowing us to identify employees in AI firms. We then merge in worker diplomas from the Education Register of Statistics Denmark, allowing us to calculate the share of college majors that work in AI firms.

A bachelor" degree in Denmark is a three year major-specific program, but most students enroll in a two-year master's degree directly following their bachelor's studies. College education is rather specialized in Denmark with 226 distinct major (`AUDD`) categories. To increase the measurement accuracy of the AI relevance scores, we pool graduating cohorts of each major into 10-year intervals. Furthermore, to reduce noise from very small cells, we require that at least 100 graduates of a college major be present in the VITA survey before assigning the major an AI relevance score. For majors that do not meet this minimum criterion, we go up one level of aggregation up (75 different categories) and assign them the more aggregated AI relevance score of this category. In calculating the AI relevances, we use the survey weights provided by Statistics Denmark to ensure that the distribution of survey firms reflects the population of firms in Denmark.

2.3 A Micro Foundation for the AI Relevance Measures

How do our AI relevance measures relate to the task-based concepts of automation emphasized in the theoretical literature on AI? Using a task-based model of production, Acemoglu and Restrepo (2018) show that AI have three distinct effects on labor demand: (i) a negative *displacement effect* in the production tasks that AI automates, (ii) a positive *reinstatement effect* in the worker tasks that are needed to facilitate AI technology, and (iii)

a positive *productivity effect* stemming from cost savings in the tasks that AI can do more efficiently than humans.

In Appendix B, we build on these insights to derive the following theoretical implications: (i) AI user relevance is the theory-consistent measure of a college major's exposure to AI's displacement effects, (ii) AI producer relevance is the theory-consistent measure of a college major's exposure to the reinstatement effects of AIs, and (iii) AI's productivity effects raise the demand for workers uniformly, and thus cancel out when we compare the relative wages for different skill groups.

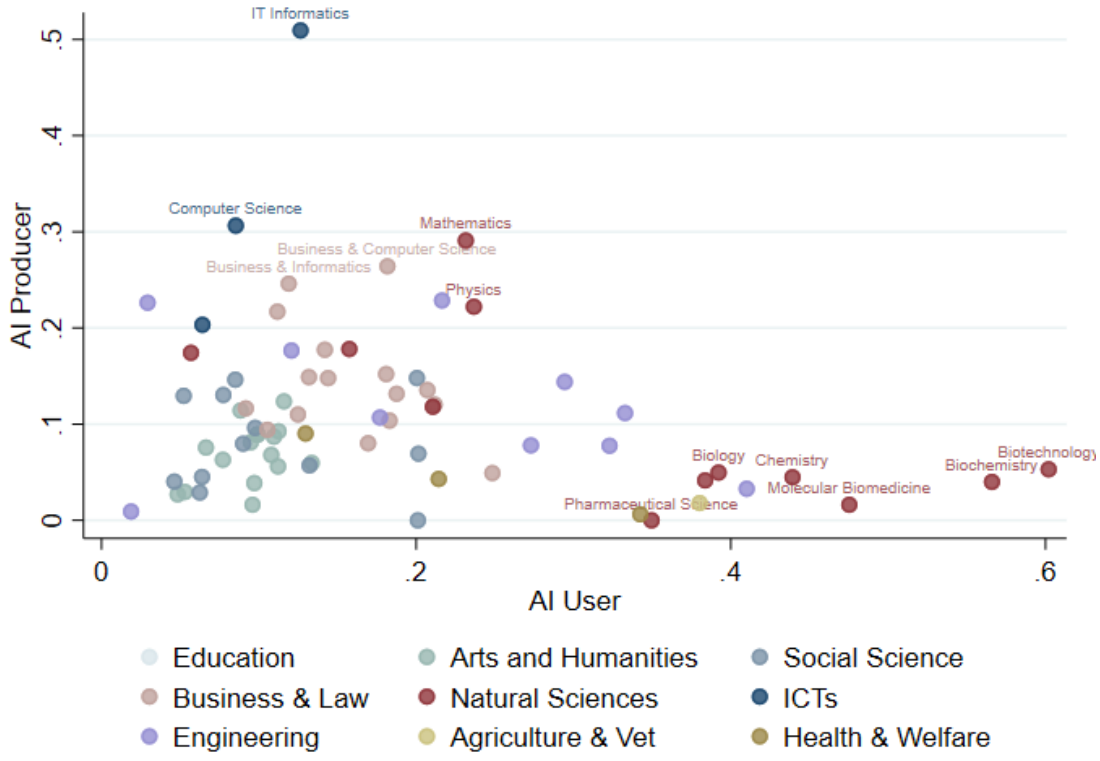
3 Facts on Artificial Intelligence and College Majors

In this section, we use the AI relevance rankings from Section 2 to document a series of facts about the college majors that are employed in AI firms. Our first key finding is that AI cuts through the group of Science, Technology, Engineering, and Mathematics (STEM) majors. While computer science and mathematics majors specialize in AI producer firms, the laboratory sciences concentrate in firms that only use AI. Our second set of facts concern the wage premiums that accrue to AI relevant majors. We find that AI producer majors earn higher wages and that these earnings premiums have increased in recent years. In contrast, we find that AI user majors do not earn higher wages, nor have experienced differential earnings growth recently.

Fact 1: Artificial Intelligence Cuts Through the STEM Degrees

Figure 2 shows a scatter plot of college majors according to their AI relevance. The figure highlights three key characteristics of AI relevant majors. First, computer science, mathematics, and physics are the most relevant majors for work with AI production. Second, biology, chemistry, and other laboratory sciences are highly concentrated in firms that only use AI. Reflecting this divide between the mathematical and laboratory sciences, engineering is also highly polarized with chemical engineers employed in AI using firms and robotics engineers employed with AI producers. Finally, the Arts and Humanities are the least relevant majors for both AI using and AI producing firms.

Figure 2: AI Relevance of College Majors

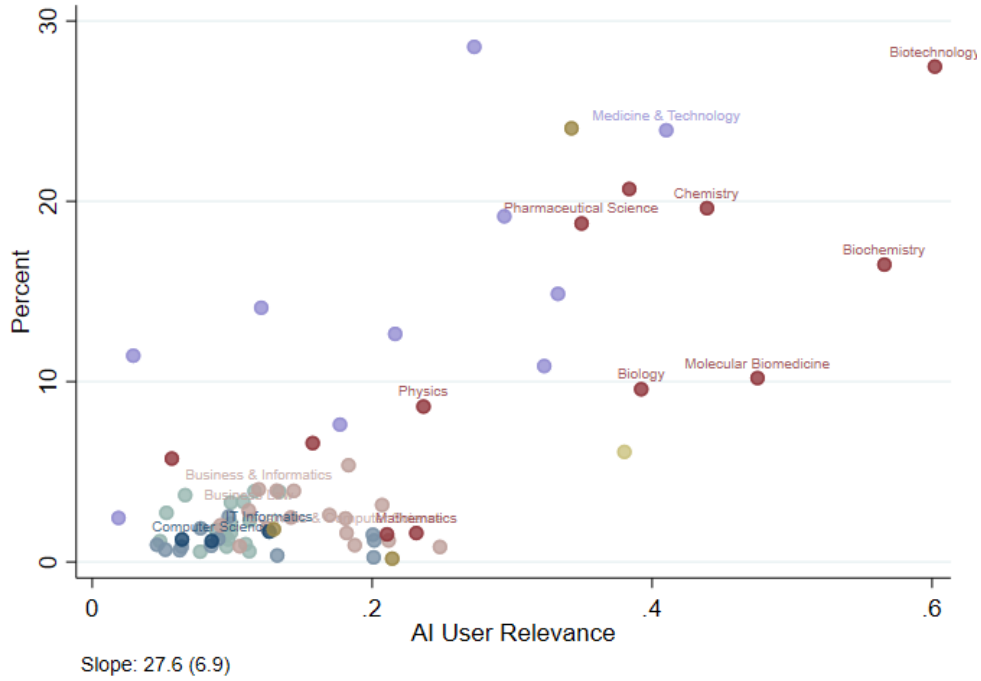


Notes: This figure plots the AI shares of college majors, defined as the fraction of college major graduates that work in AI user/producer firms. Majors are colored according to their one-digit International Standard Classification of Education (ISCED) code.

What drives the high relevance of biology and chemistry majors to AI user firms? One hypothesis is that AI has experienced technical breakthroughs in tasks that laboratory scientists traditionally carry out. Following on this idea, Webb (2020) links AI patents to job task descriptions and finds that important applications of AI relate to the automation of tasks by clinical laboratory technicians and chemical engineers. To shed light on this hypothesis, Figure 3 correlates the AI user relevance of a college major with its share of graduates that work in jobs with extreme exposure to AI applications as measured by Webb (2020). The strong positive correlation lends clear empirical support to the hypothesis that technical breakthroughs in AI applications are important determinants of the AI user relevance of different college majors. In comparison, Appendix Figure A.1 shows that AI producer relevance is not correlated with exposure to AI applications.

What characterizes the college majors that are employed in AI producer firms? Inspection of Figure 2 suggests that a strong core in computer programming or mathematics may be a necessary ingredient for AI producer relevance. However, when combined with IT capabilities, business skills seems to be a complement in AI production. For example, majors that combine business with computer science or informatics rank in the top of AI producer relevance. Interestingly, among the top 10 for AI producer relevance, business law is the only major that does not include a strong quantitative component. This could suggest that AI production requires the support services of legal experts to navigate the unchartered legal territory that surrounds AI applications.

Figure 3: Share Employed in Jobs with Extreme Exposure to AI Applications (Webb, 2020)



Notes: This figure shows the share of workers employed in occupations with extreme exposure to AI applications by AI user relevance of their college major. Extreme exposure to AI applications is defined as occupations (6-digit ISCO08) with AI exposure scores of 95/100 or above. See Webb (2020) for details about the AI exposure measure. Majors are colored according to their one-digit ISCED code; see Figure 2 for a color description. For the estimated slope coefficient, college majors are weighted by their number of graduates, and standard errors are clustered by college majors. Appendix Figure A.1 shows the corresponding figure by AI producer relevance.

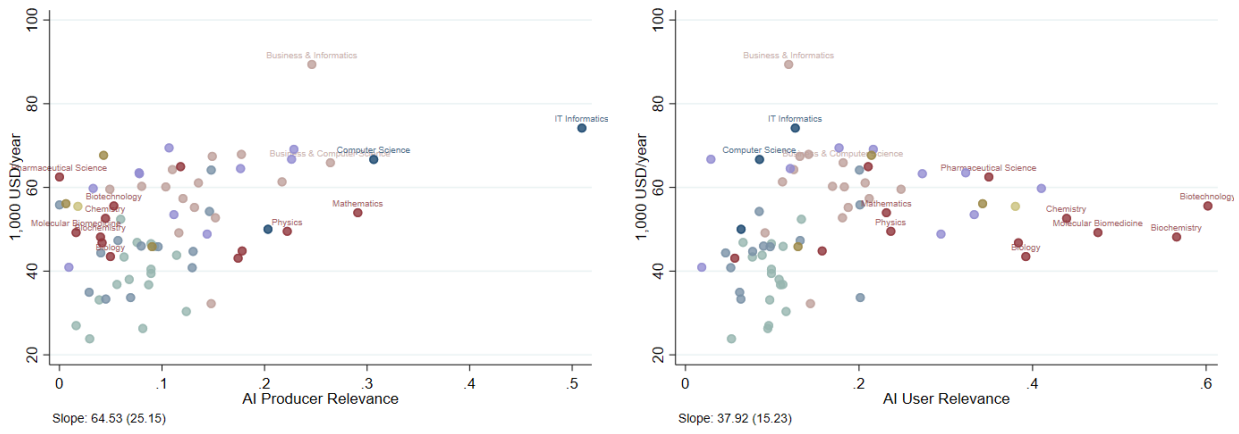
Finally, we do not find evidence that interpersonal skills alone qualify for work with AI, as otherwise often proclaimed in public debate (Trajtenberg, 2018; World Economic Forum, 2018). In fact, we find that social science and humanities are the least relevant majors for work in AI firms.

Fact 2: AI Producer Majors Earn Higher Wages, AI Users Do Not

Figure 4.(a) shows that AI producer relevance is strongly correlated with higher starting salaries when entering the labor market. A standard deviation increase in AI producer relevance (6.4 pp) is associated with about $4,100 higher annual starting salary, and the relationship is linear throughout the distribution of AI producer relevances. In contrast, Figure 4.(b) shows that starting salary does not share a similar linear relationship to AI user relevance. While extremely low AI user relevance (as for the Arts and Humanities) is associated with lower earnings, extremely high AI user relevance (as for the laboratory sciences) are not associated with correspondingly higher earnings. In fact, once we drop the Arts and Humanities from Figure 4.(b), there remains no statistically significant correlation between AI user relevance and starting salary of college graduates.

Figure 4: Starting Salaries in 2016

(a) By AI Producer Relevance of College Major (b) By AI User Relevance of College Major



Notes: This figure shows the average starting salaries of college graduates by AI relevance of their major. Starting salary is defined as salary eight years after the start of a bachelor's degree. Salary includes wage compensation for employees and net profits for self-employed. Zero incomes are included. Salary is measured in 1,000 US dollars, using an exchange rate of 5.85 Danish Kroner per US dollar. For the estimated slope coefficients, college majors are weighted by their number of graduates, and standard errors are clustered by college majors. Majors are colored according to their one-digit ISCED code; see Figure 2 for a color description.

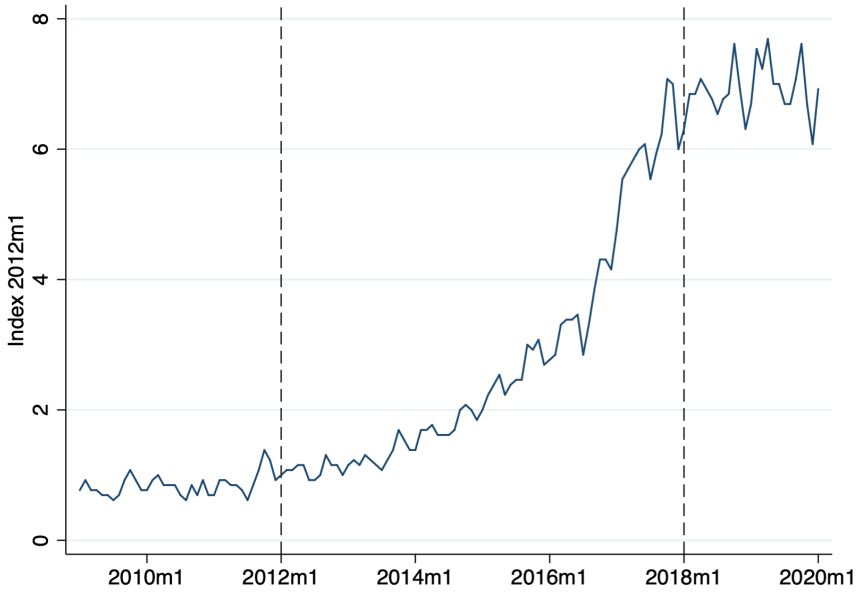
The earnings premium for AI producer skills is consistent with Alekseeva et al. (2019), who use job posting data from the United States to document large wage premiums for AI-specific programming skills such as proficiency in the machine learning packages TensorFlow and PyBrain. In Appendix Table A.1, we show that the earnings premiums of AI relevance are shared across workers with a given college major, and not confined to workers in AI firms only. This finding highlights the importance of studying AI earnings premiums throughout the economy, and not just in AI firms.

Finally, in Appendix Figure A.2, we document a sharp gender divide in the AI relevant majors, wherein men select into AI producer majors, and women select into AI user majors. For example, in the computer science degrees that are relevant for AI producers, less than ten percent of students are female. The gender divide persists within STEM degrees and suggests that an expansion of college majors according to AI relevance alone could exacerbate existing gender gaps in higher education and labor market outcomes.

Fact 3. Wage Premiums Are Rising for AI Producer Majors, Not for AI Users

To probe the timing of AI technology, Figure 5 shows that the number of Google searches on "machine learning" increased by a six-fold from 2012 to 2017.

Figure 5: Google Searches on "Machine Learning"



Notes: This figure shows the number of monthly Google searches on "machine learning" in the United States from January 2010 to January 2020. The units are normalized such that the maximum monthly search activity takes a value of 100.

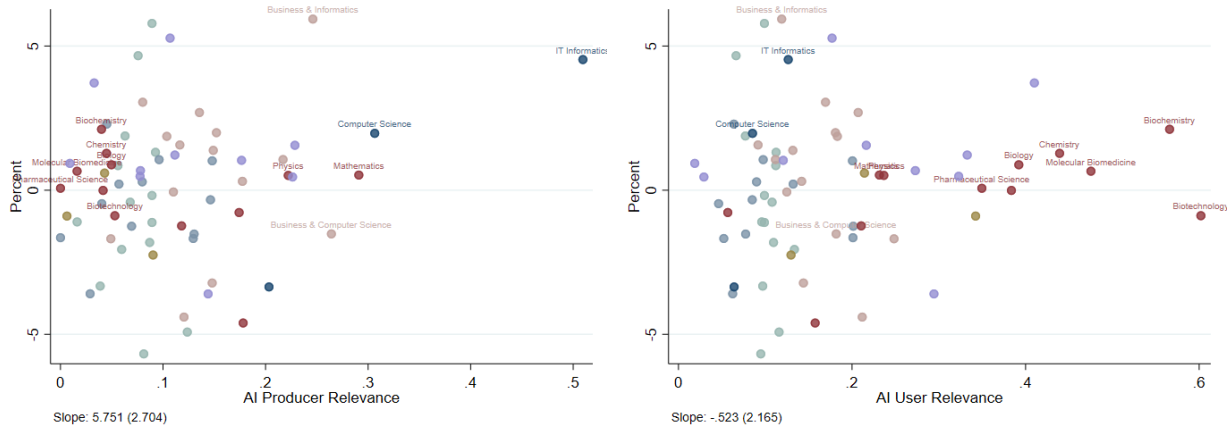
Motivated by this recent rise of interest in AI, Figure 6 plots the growth in starting salaries of college major cohorts from 2012 to 2016. Salaries are deflated using the CPI with 2015 as the base year. Panel (a) shows that AI producer majors have experienced excess earnings growth since 2012. A standard deviation higher AI producer relevance (6.4 pp) is associated with 0.4 percent points higher annual growth in starting salaries. In contrast, Panel (b) shows that AI user majors have not experienced differential earnings growth in recent years.

Thus, if AI producer relevance captures exposure to AI's task reinstatement effects as implied by our theory in Appendix B, then Figure 6.(a) suggests that these positive impacts are already beginning to show in the starting salaries of college majors. In contrast, if AI user relevance captures exposure to AI's task displacement effects, then Figure 6.(b)

shows that large negative displacement impacts of AI are yet to manifest themselves in starting salaries.

Figure 6: Annualized Growth in Starting Salaries from 2012 to 2016

(a) By AI Producer Relevance of College Major (b) By AI User Relevance of College Major



Notes: This figure shows the average annualized growth rates in starting salaries of college graduates from 2012 to 2016 by AI relevance of their major. Starting salary is defined as salary eight years after the start of a bachelor's degree. Salary includes wage compensation for employees as well as net profits for self-employed. Salaries are deflated using the CPI with 2015 as the base year. Zero incomes are included. For the estimated slope coefficients, college majors are weighted by their number of graduates, and standard errors are clustered by college majors. Majors are colored according to their one-digit ISCED code; see Figure 2 for a color description.

In Appendix Figure A.3, we replicate Figure 6 for an earlier period of years (2007-2012), showing that the differential earnings growth of AI producer majors is a recent phenomenon closely linked in time to the recent take-off in machine learning. This finding is consistent with Alekseeva et al. (2019), who document that the share of AI-skills in total job postings in the United States doubled from 2012 to 2016. Having said that, we want to emphasize that earning growth from 2012 to 2016 accounts for less than 20 percent of the observed level differences in earnings across AI producer relevance documented in Figure 4. Put differently, AI producer relevant skills have always been rewarded in the la-

bor market, but the evidence presented here suggests that AI may have exacerbated these inequalities.

We view the evidence presented in Figure 6 as consistent with Brynjolfsson et al. (2018), who observe that firms currently are undertaking significant investments into AI, but hypothesize that the productivity gains from these investments may only arrive with a delay. The investment activities are already generating positive reinstatement effects for AI producers (Figure 6.(a)), but the productivity gains of AI – an important source of potential displacement of AI user majors – are yet to manifest themselves (Figure 6.(b)).

4 Causal Evidence on Increasing AI Relevant College Skills

In Section 3 above, we documented that college majors that are relevant for AI producers earn higher wages and that these earnings premiums are on the rise. On top of these expanding wage premiums, technologists and scholars expect that the diffusion of AI will accelerate in the years to come (Brynjolfsson et al., 2018; Goldfarb et al., 2019).

How can we prepare a future generation of college students for work with AI? An often discussed policy is to adjust the admission capacities of different college majors (President's Council of Advisors on Science and Technology, 2012; Trajtenberg, 2018). However, any policy that aims to increase the AI relevance of college students by adjusting admission criteria will face at least two challenges.

First, will students, on the margin of attending an AI relevant major, complete the degree if offered a seat? Second, if the students complete the degree, will they reap the positive returns in the labor market that are suggested by the correlations in Section 3? In

this section, we study the Centralized Admission System in Denmark, where admission cutoffs allow us to provide causal evidence to each of these questions.

In Section 4.1, we first show that a substantial number of students stand on the margins of majors with widely differing AI relevances. Second, in Sections 4.3 and 4.4, we find that, if offered a spot at a more AI relevant major, these marginal students will not only complete the degree but also end up earning labor market returns that are at least as large as those suggested by the correlations.

4.1 Centralized College Admission in Denmark

The Danish university system consists of eight bachelor's granting institutions, all of which are public and use the Centralized Admission System (CAS).[1] A bachelor's degree is a three-year major-specific program, and switching between programs will, in most cases, mean reapplying through CAS. The admission capacities of different programs are centrally regulated by the Danish Ministry of Higher Education and the majority of program slots are allocated according to the high school GPA of applicants. Students can rank up to eight different major-institution combinations on the application form.

We use CAS application data from 1996 to 2007. In these years, 52 percent of applicants listed more than one priority.[2] Among these students, Figure 7 shows the range of AI relevance for their listed priorities. The takeaway from the figure is that a substantial share of students applies to majors with widely differing relevance for AI firms. In Figure

[1] All Danish university programs are free of tuition and offer the same small stipend to all students. Heinesen (2018) provides institutional details on the Centralized Admission System.

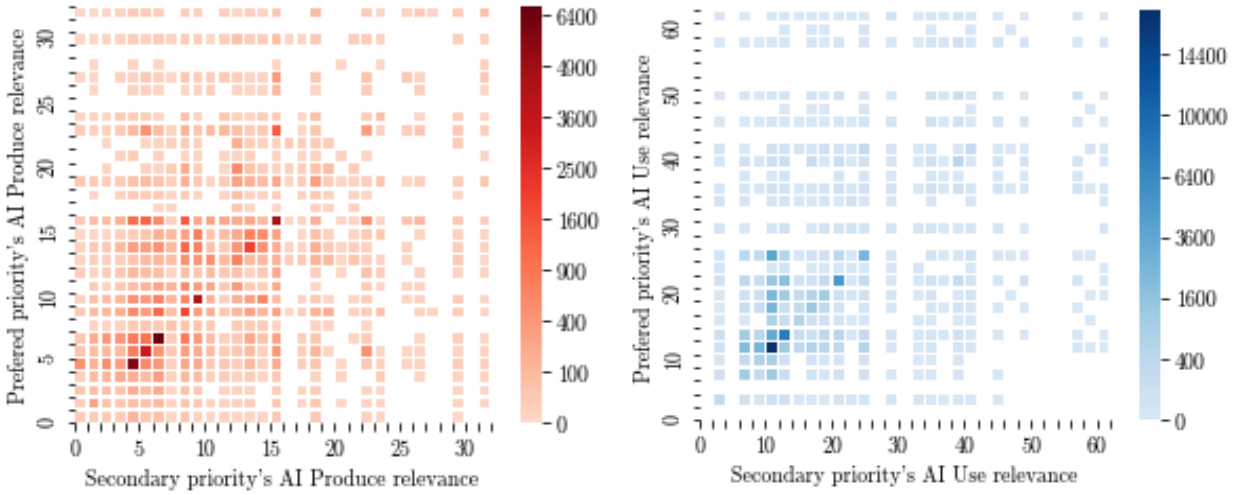
[2] Appendix Figure C.4a shows that the distributions of AI relevance in our analysis sample are representative of the population of applicants in CAS.

7a, 27 percent of students apply to programs that differ in AI producer relevance by more than five percentage points (0.9 of a standard deviation). For AI user relevance in Figure 7b, the corresponding share of students that apply for programs that differ by more than five percentage points (0.6 of a standard deviation) is 32 percent. These differences indicate that many students could have been assigned to a program of very different AI relevance had admission capacities of different programs been different.

Figure 7: AI Relevances of Priorities on College Applications

(a) By AI Producer Relevance of College Major

(b) By AI User Relevance of College Major



Notes: Color represents the total number of applications. The data is all applications from 1996-2007 with more than one priority and AI relevance measures. If there are more than two priorities, we select the two with the largest difference in AI relevance. In Panel (a), we exclude the outlier 'IT Informatics' (51 percent). Sample size is in Panel (a) is n=84,508 and in Panel (b) n=85,3687.

4.2 Regression Discontinuity Framework

This section explains how we exploit discontinuities generated by admission cutoffs to estimate causal effects of being offered a seat in a college major that is more or less AI relevant.

In the CAS, admission offers are made sequentially where the order is determined

by the applicants' GPA: the highest-ranked applicant receives an offer for her preferred program; the second-highest applicant receives an offer for her highest-ranked program among the remaining slots, and so on. This is repeated until either slots run out, or applicants run out.[3] If slots run out, a cutoff is set equal to the GPA of the last applicant accepted to the program.[4]

As students are choosing between many different majors and prioritize them differently, it is not possible to assert whether crossing an admission threshold will increase or decrease AI relevance of the offered program without information on the next-best alternative program. Hence, we follow the approach in Kirkeboen et al. (2016) and take next-best choices into account.

We form pairs of priorities using the following rules: i) The preferred program must have excess applicants and thus a GPA cutoff; the second-best may have a cutoff. ii) The cutoff of the preferred program must be higher than the second-best. iii) An applicant must not be admitted to a higher priority than the preferred program in the pair. iv) The two majors must differ in terms of AI relevance.

Table 1a illustrates an application with three priorities. The three priorities result in two potential pairs: 1st/2nd and 2nd/3rd. The applicant has a GPA of 10.0 and is accepted to her first priority, program A.[5] The relevant next-best program is her second priority, program B. This is an example where the AI relevance is highest in the preferred

[3]The allocation mechanism used in the CAS corresponds to a serial dictatorship, which is Pareto efficient and strategy proof. Serial dictatorship mechanisms are used in many college admission systems around the world, including Norway (Kirkeboen et al., 2016) and Chile (Hastings et al., 2013).

[4]The cutoffs are lax, since programs also are allowed to admit students on other criteria (*Quota Two*). In our sample, these are 18.5 percent of admissions.

[5]GPA is measured on a 10-point scale from 0 to 13 (00,03,5,6,7,8,9,10,11,13). The lowest passing GPA is 6, and 8 represents an 'average' performance.

program. Table 1b provides an example where the applicant is accepted in her second priority, program A. In this example, the more AI relevant major is ranked as next-best. The applicant is offered a seat in this AI relevant major because she does not meet the GPA cutoff of her preferred program, D.

Table 1: Application examples

(a) High AI Relevance of Preferred

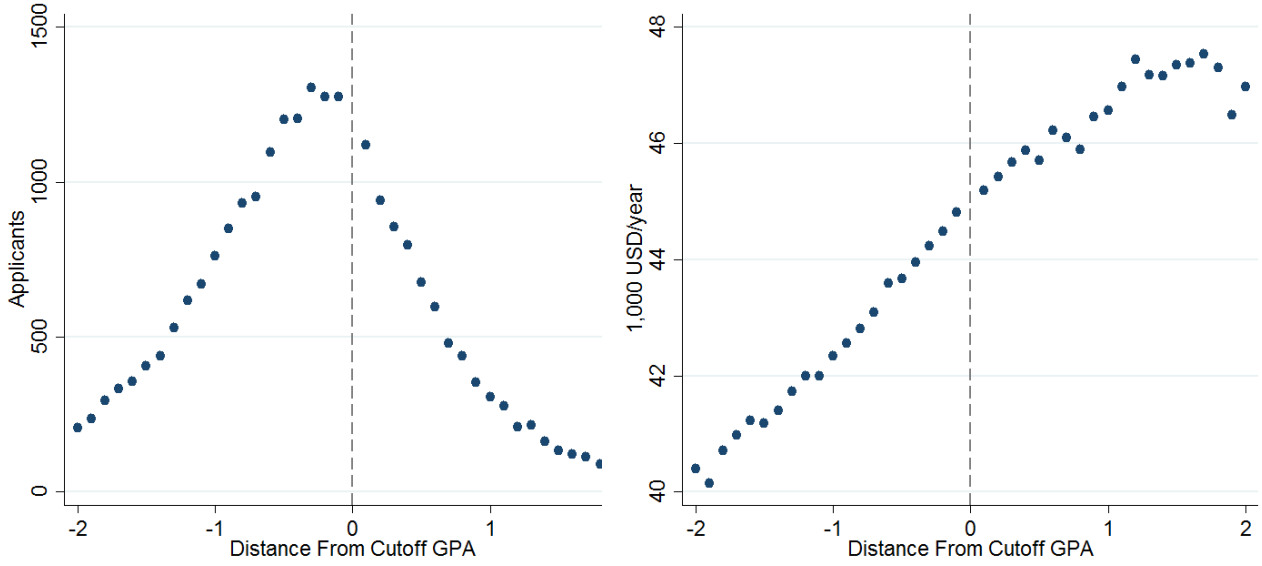
Applicant GPA = 10.0				
Priority	Program	AI Relevance	Cutoff	
1	A	0.17	9.9	Preferred
2	B	0.10	9.5	Next-best
3	C	0.15	8.0	N/A

(b) High AI Relevance of Second-best

Applicant GPA = 10.0				
Priority	Program	AI Relevance	Cutoff	
1	D	0.08	10.5	Preferred
2	A	0.17	9.9	Next-best

The admission system creates discontinuities in the likelihood of being offered a seat around the GPA cutoffs; see Appendix Figure C.3. These exact cutoffs are unpredictable at the time of application as they depend on the number of seats, applicants, and GPAs. Our identifying assumptions are standard for RD designs. We assume that there are no other changes occurring at the thresholds that could confound our estimates; see Lee (2008) and Lee and Lemieux (2010) for a thorough treatment. Figure 8 represents robustness checks, showing no manipulation in the density of the running variable (GPA minus cutoff) and smooth behavior of predicted earnings near the threshold.

Figure 8: Manipulation Around the Cutoff



(a) Applicant Density (b) Predicted Earnings

Notes: We pool all admission cutoffs and normalize so that zero on the x-axis represents the admission cutoff. The sample is all priorities of programs with cutoffs. (b) Variables used in prediction: age at the time of application, gender, and high school GPA.

Our main estimating equation is

$$y_i = \alpha + \beta Z_i + \alpha_1 d_i + \alpha_2 d_i^2 + (\alpha_3 d_i + \alpha_4 d_i^2) \times Z_i + \eta X_i + \mu_i \quad (1)$$

where y_i represents an outcome of applicant i, such as degree attainment or earnings. d_i denotes the distance from applicant GPA to the program cutoff. The indicator variable Z_i equals one if the applicant's GPA is above the program cutoff and zero otherwise. Our parameter of interest is the causal effect of admission threshold crossing, β. We use a bandwidth of two grade points, and a second-order polynomial interacted with Z to allow for different slopes on each side of the cutoff. We also include gender and age as a vector of predetermined control variables, X. In practice, these control variables have

very little effect on our RD estimates, and mainly serve to improve precision.

4.3 AI Relevance of Completed College Major

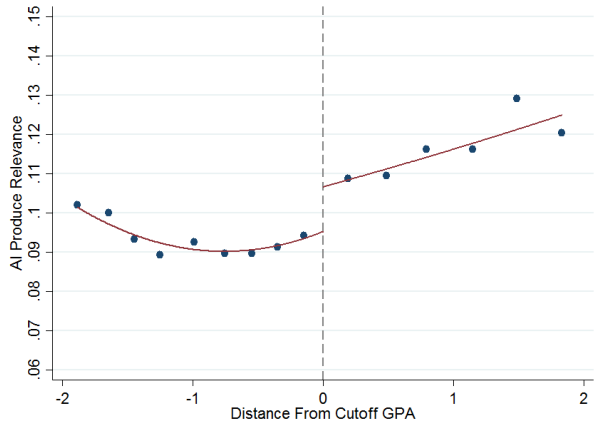
Figure 9 shows how crossing an admission threshold impact the AI relevance of the completed major eight years later. The discontinuities around the thresholds correspond to the RD estimates of β in Equation (1), which we also report in Table 2.(b).

Panels (a) and (b) show the impact on AI producer relevance of completed major. Crossing an admission threshold to a more (less) AI producer relevant program raises (lowers) the AI relevance of a student's completed degree by 1.1 (1.9) percentage points on average. As the discontinuities in offered programs are 2.6 and 2.9 percentage points (Panel (a) of Table 2), this implies a passthrough rate from offered seat to completed degree of around 40 to 60 percent. Panels (c) and (d) show the threshold crossing effects on AI user relevance, which imply similar passthrough rates of around 30 to 70 percent.

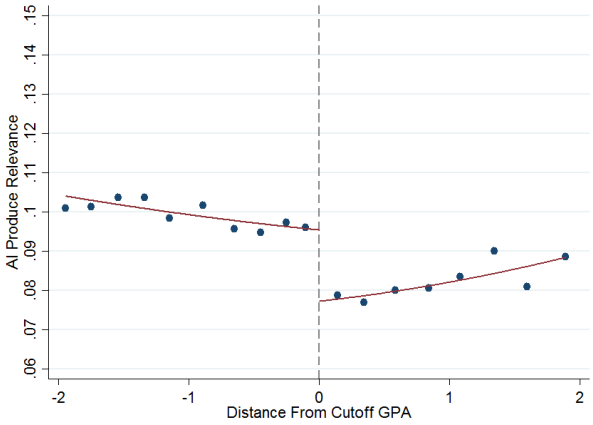
To be able to calculate an average realized AI relevance, Figure 9 focuses on applicants who have completed some bachelor's degree eight years later. Thus, a natural follow-up question is whether higher AI relevance comes the price of lower bachelor's completion rate. In Appendix Figure C.5, we show that there is no discontinuity in completion rates, eight years later, when crossing the admission threshold to a more AI relevant degree. However, we estimate an increase in completion rates of around five percentage points when crossing the threshold to a less AI relevant major. Put differently, it is only students who do not want a more AI relevant major whose completion rates suffer when pushed toward these AI relevant programs.

Figure 9: AI Relevance of Completed Degree and Threshold Crossing

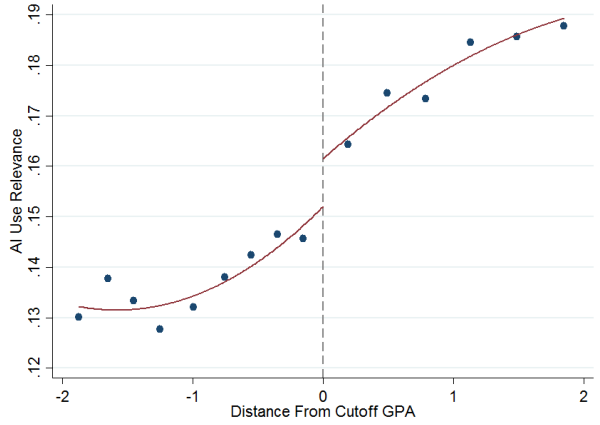
(a) Preferred Degree Is More AI Producer Relevant



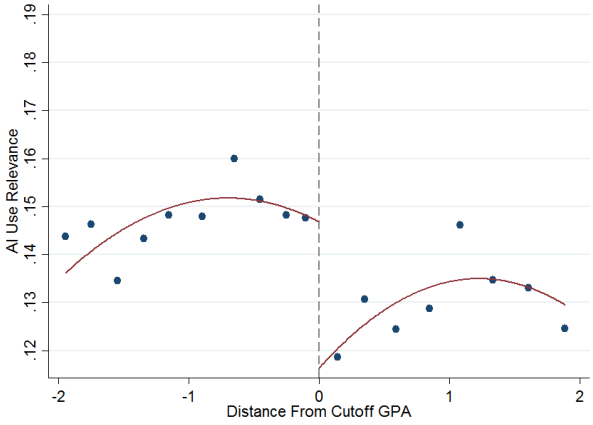
(b) Next-Best Is More AI Producer Relevant



(c) Preferred Degree Is More AI User Relevant



(d) Next-Best Is More AI User Relevant



Notes: The figures plot the mean of AI relevance of completed degrees, eight years after application, within bins of GPA minus cutoff, and fit estimated quadratics using all the underlying data. Subsample of applicants who eight years later has completed a bachelor's degree categorized with the AI relevance measure. Excluding applicants exactly on the cutoff.

The estimated passthrough rates from offered seat to completed degree are centered around 50 percent,[6] which is sizable and suggests that adjusting admission cutoffs for different majors may be an effective way to increase the AI relevance of college grad-

[6]In the Danish admission system, several institutional features will tend to make the passthrough rate low. For example, applicants are free to reapply again next year or apply to their preferred program via the quota 2 system (see footnote 4)

uates.[7] These findings are noteworthy, given that we in Section 3 documented that AI producer relevance requires hard skills that could make it difficult for a marginal student to complete a program.

4.4 AI Relevance and Earnings

Figure 9 shows how crossing admission thresholds impact early-career salaries of college graduates. Panel (a) of the figure shows that crossing the threshold to a more AI producer relevant major leads to a statistically significant increase in yearly earnings of $3,770. In comparison, Panel (b) shows that crossing the threshold to a less AI producer relevant degree leads to an insignificant drop in yearly earnings of $1,692. In contrast, in Panels (c) and (d), we estimate that the causal returns to higher AI user relevance is a precisely estimated zero. Table 2.(c) reports the corresponding regression estimates.

Table 2 shows a noteworthy asymmetry in the effects of AI producer relevance, whereby students have higher bachelor's completion rates and receive larger earnings gains of AI producer relevance if they also prefer the AI relevant degree.[8] The asymmetry in completion rates can only partly explain the asymmetry in earnings, which persists even when we condition on students who complete their bachelor's studies. This estimated asymmetry in earnings effects of AI producer relevance is consistent with students sorting on individual comparative advantages, as documented in the Norwegian setting by Kirke-

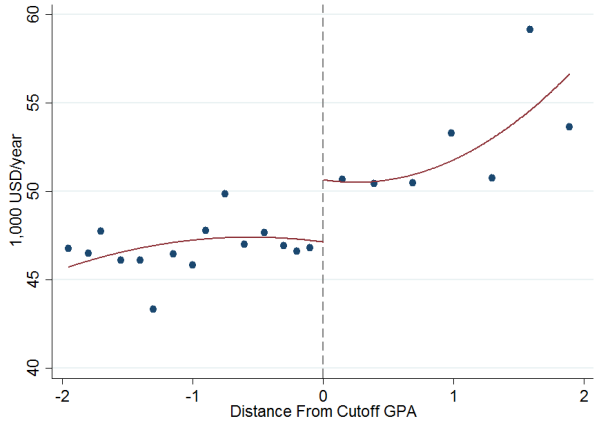
[7] To evaluate any specific policy intervention that adjusts admission cutoffs, for example, by controlling the admission capacities of different majors, one would need to take into account where in the AI relevance distribution the marginal students are drawn in from. This would require a reweighting of the underlying marginal treatment effects relative to the local average treatment effects identified by our estimation strategy. See Mogstad et al. (2018) for a formal framework for using instrumental variables to draw inference about policy-relevant treatment parameters.

[8] Figure C.5 showed that the asymmetry in completion rate only applies to students who do not prefer the AI relevant degree

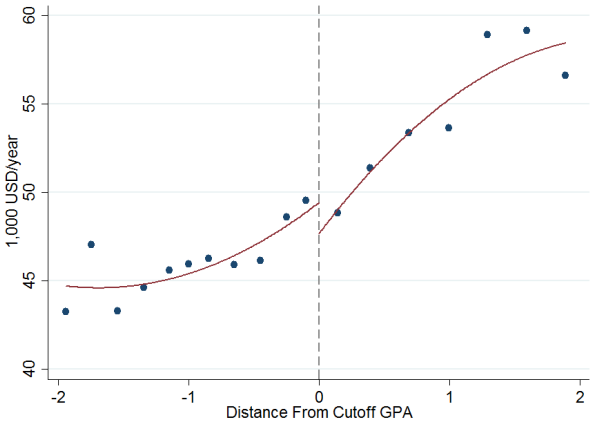
boen et al. (2016). If applicants sort according to comparative advantages, it can explain why students experience larger gains from AI producer relevance when they also rank these programs highest. From a policy perspective, these findings suggest that it may be more effective to expand AI relevant programs to students who prioritize them rather than to students who list them as an alternative.

Figure 10: Earnings and Threshold Crossing

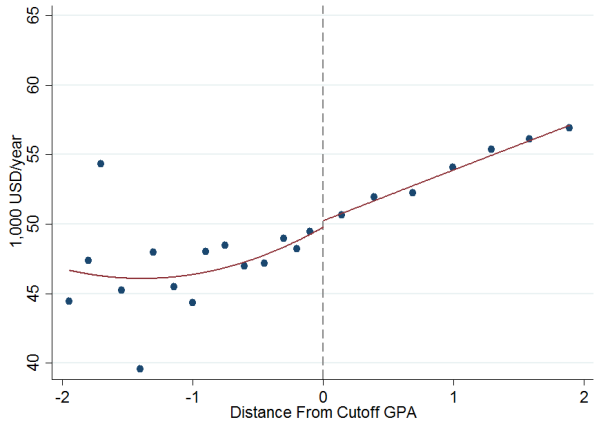
(a) Preferred Degree Is More AI Producer Relevant



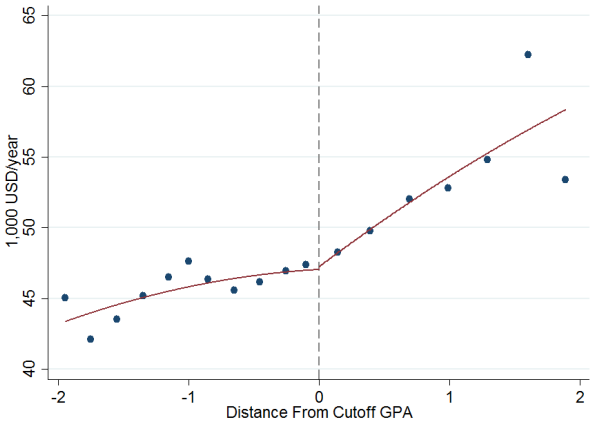
(b) Next-best Is More AI Producer Relevant



(c) Preferred Degree Is More AI User Relevant



(d) Next-Best Is More AI User Relevant



Notes: The figures plot the mean of earnings within bins of GPA minus cutoff, and fit estimated quadratics using all the underlying data. Excluding applicants exactly on the cutoff and restricting to earnings above zero. Earnings are the median of wage income plus self-employed business income in the years seven to nine years after application. Income variables are deflated using the CPI with 2015 as the base year.

To put the effects of AI producer relevance into perspective, in Appendix Table C.2, we compare them to the OLS earnings correlations estimated throughout the sample, and not just around the admission thresholds. For students who rank their AI relevant major as preferred, the causal effect of AI producer relevance is noticeable larger than the correlation. For students who rank the AI relevant major as next-best, the causal effects of AI producer relevance are similar to those suggested by the correlation. These findings suggest that the earnings premiums of AI producer relevance documented in Section 3 reflect causal effects of obtaining AI relevant college skills. The findings highlight that college education may be a powerful tool to prepare future generations for work with AI.

In summary, the findings in this section suggest that AI producer relevance may be a valuable target for college admission policies. However, how do the AI producer earnings effects compare to other targets that policymakers could follow when deciding which programs to allocate slots? A commonly used target in the Danish context is to rank programs according to their early-career average salaries (Kvalitetsudvalget, 2014). To compare our estimates to this benchmark, Appendix C.3 reruns our RD analysis, now ranking the majors according to their average starting salaries. Interestingly, we find that pushing students according to AI producer relevance yields larger causal earnings returns than simply pushing them toward degrees with higher average salaries. This suggests that it indeed is possible to acquire AI producer relevant skills in college, and that the earnings premiums related to these college skills are less driven by selection than the overall earnings differences between majors. This is a noteworthy finding, as AI producer relevance require quantitative skills that could be difficult to acquire late in life without prior training or inclination.

Table 2: Threshold Crossing Effects

(a) AI Relevance of Offered Degree

	High AI Preferred		Low AI Preferred	
	Producer Relevance	User Relevance	Producer Relevance	User Relevance
	(1)	(2)	(3)	(4)
RD Estimate	0.0261	0.0294	-0.0325	-0.0477
	(0.0029)	(0.0042)	(0.0026)	(0.0050)
Mean below cutoff	0.066	0.133	0.0936	0.133
Observations	8550	11665	12328	9221

(b) AI Relevance of Completed Degree

	High AI Preferred		Low AI Preferred	
	Producer Relevance	User Relevance	Producer Relevance	User Relevance
	(1)	(2)	(3)	(4)
RD Estimate	0.0112	0.00915	-0.0182	-0.0321
	(0.0035)	(0.0048)	(0.0030)	(0.0056)
Mean below cutoff	0.0922	0.140	0.0981	0.149
Observations	5990	8100	8633	6530

(c) Earnings

	High AI Preferred		Low AI Preferred	
	Producer Relevance	User Relevance	Producer Relevance	User Relevance
	(1)	(2)	(3)	(4)
RD Estimate	3770.4	609.84	-1592.3	357.93
	(1593.6)	(1378.9)	(1378.9)	(1494.2)
Mean below cutoff	47151.6	47334.7	46517.6	46062.3
Observations	8550	11665	12328	9221

Notes: Estimation of equation (1) with gender and age as control variables. (b) is a subsample of applicants who, eight years later, has completed a bachelor's degree categorized with the AI relevance measure. Excluding applicants exactly on the cutoff. Bandwidth is +/- 2.0 grade points.

5 Concluding Remarks

This paper makes two contributions to our understanding of how artificial intelligence relates to the returns of college majors. First, we use new micro data from Denmark to rank college majors according to whether their graduates work in AI firms. Second, we provide causal estimates on the effects of pushing students toward more AI relevant degrees.

A takeaway from our study is that traditional fields of study is an imperfect taxonomy for understanding AI relevance, and that AI cuts through the group of STEM degrees in particular. In further work, we welcome efforts to understand better what exact course contents are needed to prepare college students for work with AI.

An upshot of our regression discontinuity analysis is that adjusting admission cutoffs may be a powerful policy margin to push students toward more AI relevant skills. Our estimated causal effects inform policies that aim to increase the AI relevance of college students by adjusting admission capacities at different college majors, and we welcome further efforts to specify and evaluate such policy interventions.

Finally, while we estimated significant positive returns to the college majors that are relevant for AI producers, we also documented that men currently are heavily overrepresented in these same programs. To ensure that the fruits of AI are shared equally among men and women, we encourage further efforts to increase the female representation in the AI producer majors.

References

Acemoglu, D. and P. Restrepo (2018). Artificial Intelligence, Automation and Work.

Acemoglu, D. and P. Restrepo (2020, jun). Robots and Jobs: Evidence From US Labor Markets. *Journal of Political Economy*.

Alekseeva, L., J. Azar, M. Gine, S. Samila, and B. Taska (2019). The Demand for AI Skills in the Labor Market. *SSRN Working Paper*, 1–38.

Brynjolfsson, E., D. Rock, and C. Syverson (2018). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. *NBER Working Paper 24001*.

Goldfarb, A., B. Taska, and F. Teodoridis (2019). Could Machine Learning Be a General-Purpose Technology? Evidence from Online Job Postings.

Goldin, C. and L. F. Katz (1998). The Origins of Technology-Skill Complementarity. Technical Report 3.

Hastings, J. S., C. a. Nielson, and S. D. Zimmerman (2013). Are Some Degrees Worth More Than Others? Evidence From College Admission Cutoffs in Chile. *National Bureau of Economic Research Working Paper Series 19241*, 1–50.

Heinesen, E. (2018). Admission to Higher Education Programmes and Student Educational Outcomes and Earnings: Evidence From Denmark. *Economics of Education Review 63*, 1–19.

Kirkeboen, L., E. Leuven, and M. Mogstad (2016). Field of Study, Earnings and Self-selfselection. *The Quarterly Journal of Economics 131*(3), 1057–1111.

Kvalitetsudvalget (2014). Nye Veje Og Høje Mål. Technical report.

Lee, D. S. (2008). Randomized Experiments From Non-Random Selection in U.S. House Elections. *Journal of Econometrics 142*, 675–697.

Lee, D. S. and T. Lemieux (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature 48*(2), 281–355.

Mogstad, M., A. Santos, and A. Torgovitsky (2018). Using Instrumental Variables for Inference About Policy Relevant Treatment Parameters. *Econometrica 86*(5), 1589–1619.

President's Council of Advisors on Science and Technology (2012). Report To the President Engage To Excel: Producing One Million Additional College Gr Aduates W Ith Degrees in Science, Technology, Engineer Ing, and Mathematics. Technical report, Washington D.C.

Trajtenberg, M. (2018). AI As The Next GPT: A Political-Economy Perspective. *The Economics of Artificial Intelligence: An Agenda*, 175–186.

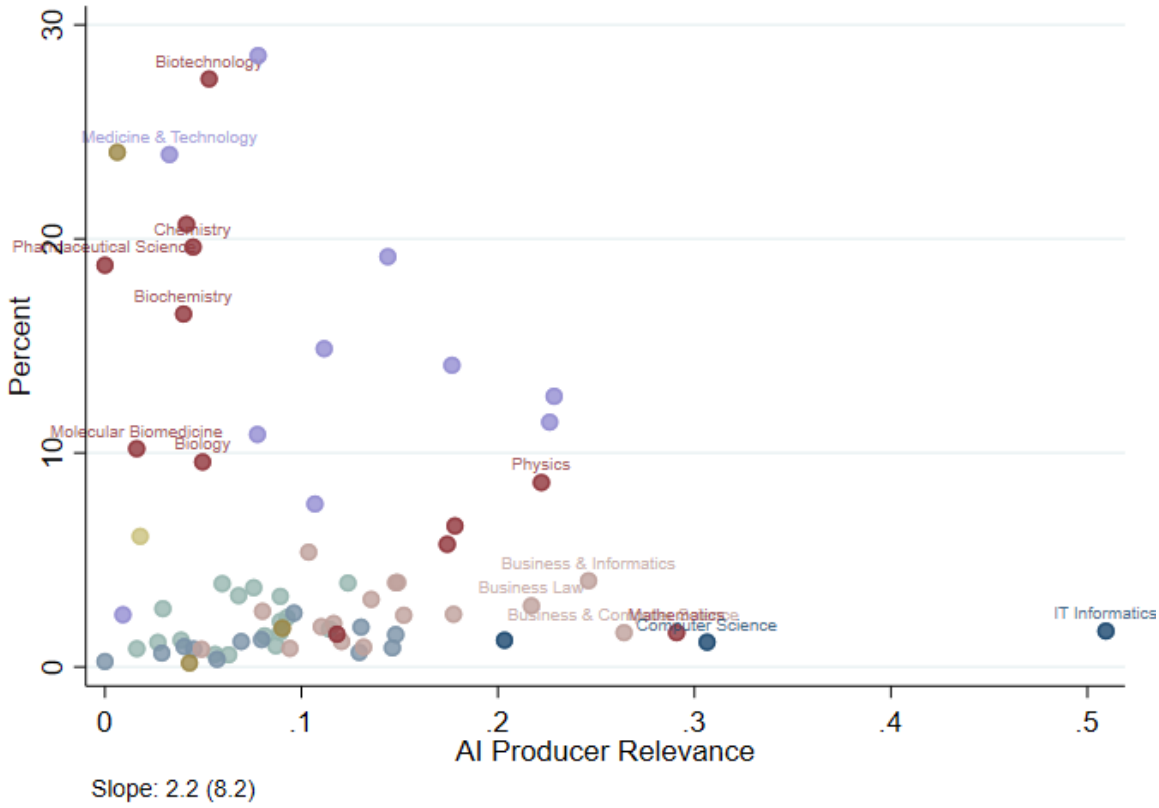
Warzynski, F. (2020). AI Adoption in Denmark.

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market.

World Economic Forum (2018). 10 Skills You'll Need to Survive the Rise of Automation — World Economic Forum.

A Facts on Artificial Intelligence and College Majors

Figure A.1: Share Employed in Jobs with Extreme Exposure to AI Adoption



Notes: This figure shows the share of workers employed in occupations with extreme exposure to AI technologies by AI producer relevance of their college major. Extreme exposure to AI adoption is defined as occupations (4-digit ISCO88) with AI exposure scores of 95/100 or above. See Webb (2020) for details about the AI exposure measure. Figure 3 shows the corresponding figure by AI user relevance.

Table A.1: Earnings Premiums of AI Relevance and AI Firms

(a) Starting Salary in 2016 and AI Producer Relevance

	(1)	(2)	(3)	(4)
AI Producer Relevance	52.9		44.7	47.3
	(16.1)		(15.1)	(18.6)
AI Producer Firm		9.2	6.6	8.4
		(1.7)	(1.6)	(4.7)
AI Producer Relevance × AI Producer Firm				-11.4
				(23.0)
Observations	2,424	2,424	2,424	2,424
R^2	0.028	0.018	0.037	0.037

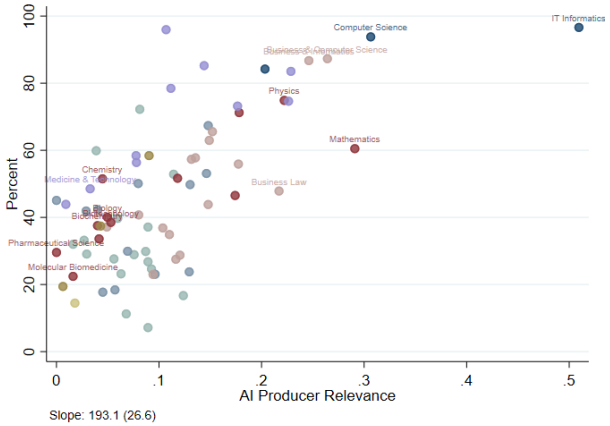
(b) Starting Salary in 2016 and AI User Relevance

	(1)	(2)	(3)	(4)
AI User Relevance	31.6		18.6	23.2
	(10.8)		(10.3)	(11.3)
AI User Firm		11.7	10.1	12.7
		(2.0)	(1.8)	(3.2)
AI User Relevance × AI User Firm				-12.9
				(12.5)
Observations	2,424	2,424	2,424	2,424
R^2	0.016	0.033	0.038	0.038

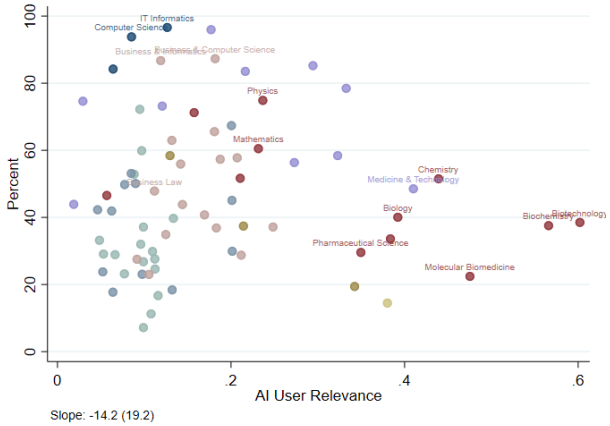
Notes: This table shows estimates from regressions of starting salaries on AI relevance of college majors (Row 1), an indicator for working in an AI firm (row 2), and the interaction of AI relevance and working in an AI firms (row 3). Starting salary is defined as salary eight years after the start of a bachelor's degree. Salary includes wage compensation for employees and net profits for self-employed. Zero incomes are included. Standard errors are clustered at the college major level.

Figure A.2: Male Share and AI Relevance of College Majors

(a) AI Producer Relevance



(b) AI User Relevance



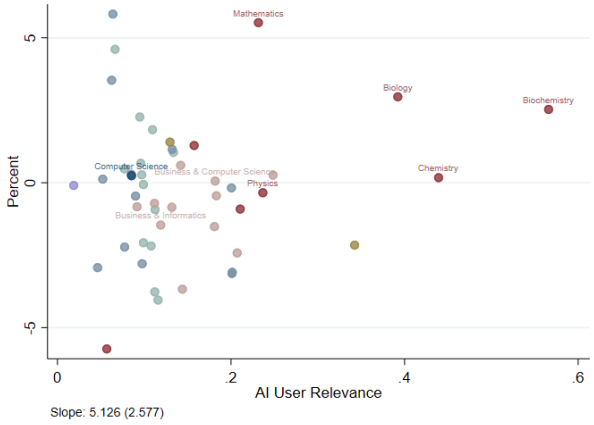
Notes: This figure shows the share of males by AI relevance of college majors. For the estimated slope coefficients, college majors are weighted by their number of graduates, and standard errors are clustered by college majors. Majors are colored according to their one-digit ISCED code; see Figure 2 for a color description.

Figure A.3: Growth in Starting Salaries 2007-2012 by AI Relevance of College Major

(a) AI Producer Relevance



(b) AI User Relevance



Notes: This figure shows the average annualized growth rates in starting salaries of college graduates from 2007 to 2012 by AI relevance of their major. Salary includes wage compensation for employees as well as net profits for self-employed. Zero incomes are included. Salaries are deflated using the CPI with 2015 as the base year. For the estimated slope coefficients, college majors are weighted by their number of graduates, and standard errors are clustered by college majors. Majors are colored according to their one-digit ISCED code; see Figure 2 for a color description.

B A Micro Foundation for the AI Relevance Measures

This section provides a task-based micro foundation for the AI relevance measures proposed in Section 2. In our stylized model, the economy consists of three types of firms: AI users (u), AI producers (p), and non-AI firms (n).

AI users produce by combining a set of production tasks i

$$Y_u = \exp\left(\int_0^1 X_u(i)di\right) \tag{2}$$

Production tasks can be carried out by either labor or AI. We follow Acemoglu and Restrepo (2020) and assume that AI can automate a fraction θ of tasks

$$X_u(i) = \begin{cases} L_u(i) + \gamma(i) \times \text{AI}(i) & \text{if } i \leq \theta \\ L_u(i) & \text{if } i > \theta \end{cases} \tag{3}$$

where $\gamma(i)$ denotes the relative productivity of AI in task i, and L_u is the labor input bundle of AI user firms

$$L_u(i) = \prod_{s \in \mathcal{S}} L_{us}^{\alpha_{us}}(i) \tag{4}$$

AI is developed, implemented, and maintained by AI producer firms, who thus reinstate

the demand for workers that are employed in these firms

$$\text{AI}(i) = Y_p(i) = \prod_{s \in \mathcal{S}} L_{ps}^{\alpha_{ps}}(i) \tag{5}$$

Our incorporation of AI production is similar to Goldin and Katz (1998), who specify that the adoption of advanced machinery requires specialized inputs from skilled machinists.

Final output in the economy is a bundle of output from AI-user and non-AI firms

$$Y = Y_u^{\beta} Y_n^{1-\beta}$$

where non-AI firms operate a Cobb-Douglas production technology with skill shares α_{ns}.

Product and labor markets are competitive with real wages W_s. Following Acemoglu and Restrepo (2020), we assume that AI is cost-effective in the tasks that are technologically feasible to automate.[9]

From cost minimization, the demand for workers of skill s is

$$W_s L_s = Y\left((1-\beta)\alpha_{ns} + \beta((1-\theta)\alpha_{us} + \theta\alpha_{ps})\right)$$

Finally, also following Acemoglu and Restrepo (2020), we model technical progress in

[9]Formally, for AI to be cost-effective, we assume that

$$\prod_{s \in \mathcal{S}} \left(\frac{W_s}{\gamma(i)\alpha_{ps}}\right)^{\alpha_{ps}} \leq \prod_{s \in \mathcal{S}} \left(\frac{W_s}{\alpha_{us}}\right)^{\alpha_{us}}, \quad \forall i \leq \theta \tag{6}$$

AI as a shift in θ. The impact of AI on the demand for workers of skill s is then

$$d\log(W_s L_s) = \underbrace{d\log(Y)}_{\substack{\text{Productivity}\\ \text{Effect}}} - \underbrace{R_{us} \times \frac{d\theta}{1-\theta}}_{\substack{\text{Displacement}\\ \text{Effect}}} + \underbrace{R_{ps} \times \frac{d\theta}{\theta}}_{\substack{\text{Reinstatement}\\ \text{Effect}}} \tag{7}$$

where R_{us} and R_{ps} denote the AI relevance measures proposed in Section 2. To recall, AI relevance is defined as the share of workers of skill s who are employed in AI firms.[10]

Equation (7) shows that AI impacts the demand for worker skills through three channels: a positive productivity effect, a negative displacement effect, and a positive reinstatement effect. The equation delivers three key insights about how AI affects the demand for skills. First, the productivity effect, which is positive because AI is cost-effective at the margin, is shared uniformly across workers of different skills. This effect will lift all boats and cancels out when we compare the relative demands for different skill groups. Second, AI user relevance R_{us} is the theory-consistent measure of a skill group's exposure to the negative displacement effects of AI automation, $-\frac{d\theta}{1-\theta}$. Finally, AI producer relevance R_{ps} is the theory-consistent exposure measure to the positive reinstatement effects coming from AI production, $\frac{d\theta}{\theta}$.

[10] According to our theory, the AI relevance measures have the following structural expressions

$$R_{ps} = \frac{L_{ps}}{L_s} = \frac{\beta\theta\alpha_{ps}}{(1-\beta)\alpha_{ns} + \beta((1-\theta)\alpha_{us} + \theta\alpha_{ps})}$$
$$R_{us} = \frac{L_{us}}{L_s} = \frac{\beta(1-\theta)\alpha_{us}}{(1-\beta)\alpha_{ns} + \beta((1-\theta)\alpha_{us} + \theta\alpha_{ps})}$$

C The Danish Centralized Admission System (CAS)

C.1 CAS Data

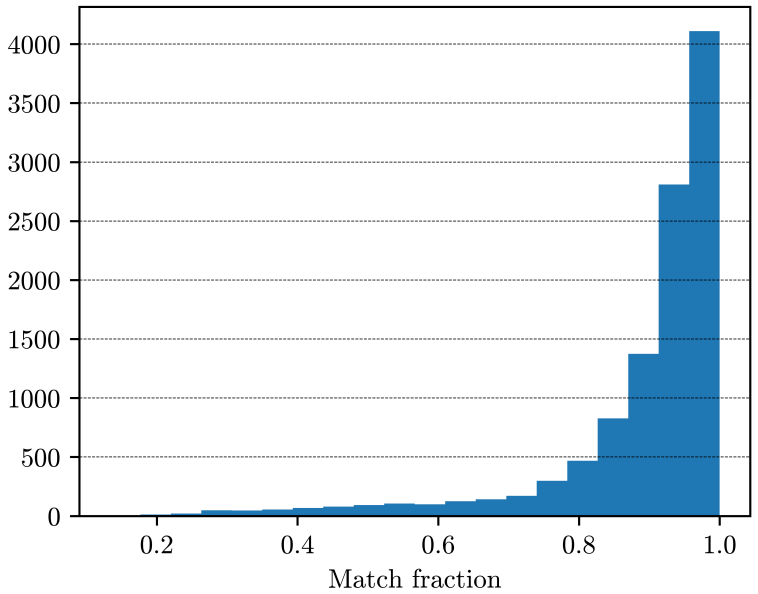
We use CAS data from 1996-2007 merged with register data from Statistics Denmark (DST). Programs in nursing, teaching, child care, and social work are not offered by universities, and we do not consider these. In CAS, each program has a *kotnr*-code. These are not easily converted to the field codes of DST (*audd/udd*). Furthermore, *kotnr* is not stable over time and only unique within each application year. Hence we crosswalk from *kotnr* to *audd* with the following procedure. We compare the *kotnr* with the enrolled programs in DST's education register (*KOTRE*) for all accepted applicants each year. A student does not necessarily appear in the offered program in the DST register as it tracks actual enrollment in October. Not all students accept the offered seat, and some are able to enroll in other programs with empty slots. Hence, we crosswalk to the most prevalent *audd* code. We only crosswalk *kotnr*s when there is: i) At least ten applicants ii) At least 50 percent of the students are enrolled in the most prevalent *audd* code. Figure C.1 shows the distribution of match quality; the fraction of students with the most prevalent code.

Table C.1: Applications in CAS

Complete applications 1996-2007	274,067
High School GPA available	233,502
Income data available	229,811
More than one priority	119,715
At least one priority with cutoff	102,599
Forming pairs within applications:	
Relevant pairs	56,941
Both priorities crosswalked and classified in VITA firm survey	37,327
Final samples:	
i) Priorities differ in AI produce relevance	26,842
ii) Priorities differ in AI use relevance	26,850

Notes: Complete applications are applications with at least one priority crosswalked to a university program and does not include students with standby priority from the year before (as these have priority regardless of cutoffs). Only applicants with high school GPA and income data in DST's registers. This implies that we only use applicants with a Danish High School exam, who are living in the country seven to nine years later. Whenever a program has fewer applicants than slots; there is no cutoff. The amount of relevant pairs is the result of the first three rules described in section 4. The loss of observations in the crosswalk to AI relevance step is mainly due to certain *audd*-codes not being well represented in the VITA firm survey.

Figure C.1: Crosswalk Match Quality Distribution

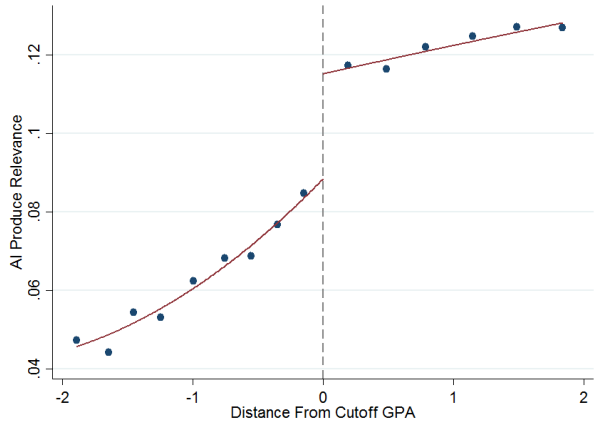


Notes: The figure shows a histogram fraction of students who are enrolled in the most typical field code (*audd*) by CAS program code (*kotnr*). We only crosswalk codes with a better match fraction than 0.5.

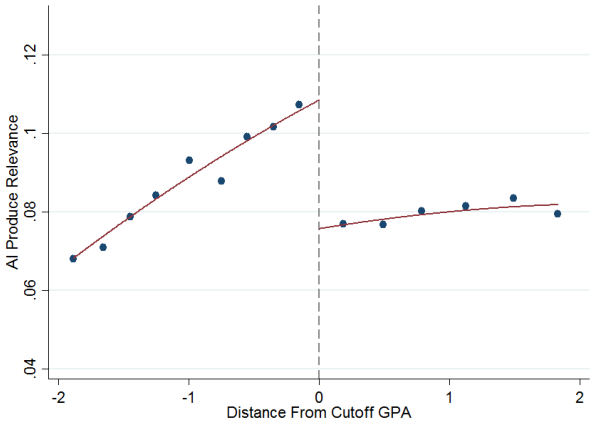
C.2 Figures and Tables

Figure C.2: Offered Program and Threshold Crossing

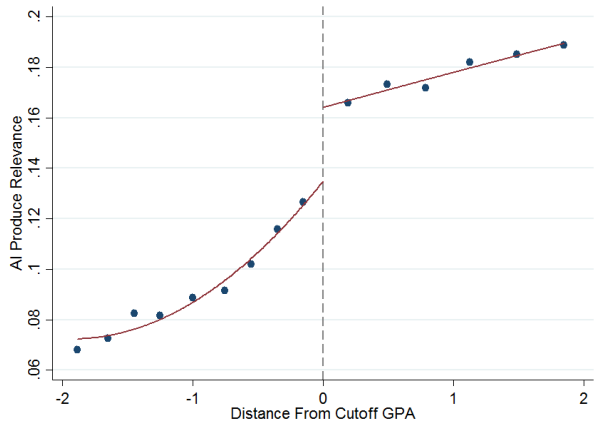
(a) Preferred Degree Is AI Producer Relevant



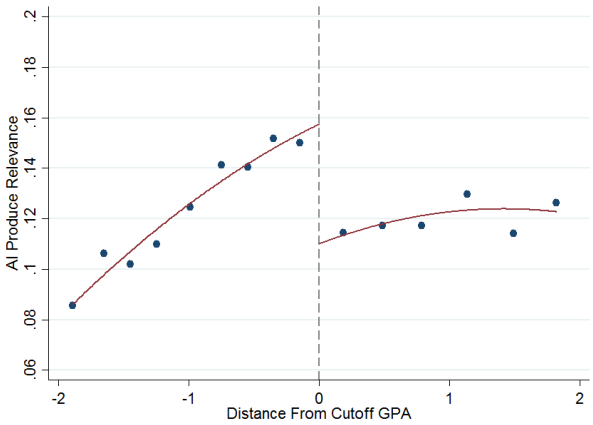
(b) Next-Best Is AI Producer Relevant



(c) Preferred Degree Is AI User Relevant

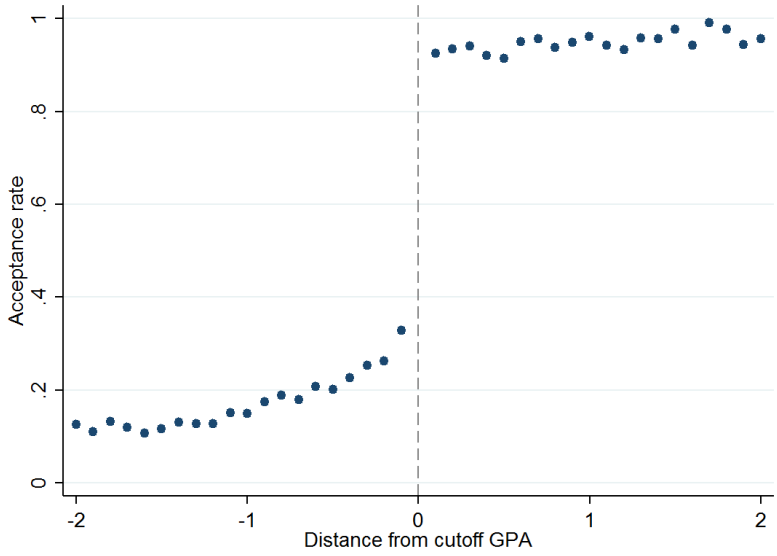


(d) Next-Best Is AI User Relevant



Notes: The figures plot the mean of AI relevance of offered degrees from the CAS system within bins of GPA minus cutoff, and fit estimated quadratics using all the underlying data. Excluding applicants exactly on the cutoff.

Figure C.3: Distance to Cutoff and Acceptance

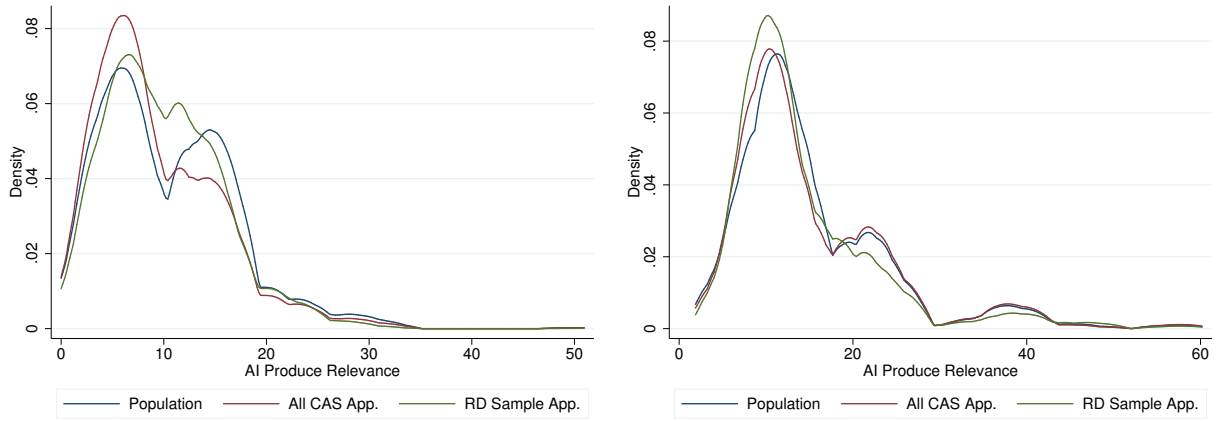


Notes: The figure shows the mean of an indicator for acceptance binned by tenths of a grade point on the x-axis. If students also apply for a standby seat (granted enrollment latest one year later), the standby cutoff is used. The accepted applicants below the cutoff are Quota Two. Applicants not accepted above the cutoff are due to lack of program-specific prerequisites. We are forced to exclude applicants exactly on the cutoff, as we only have GPAs and cutoffs with the first one decimal place in our data. CAS uses a lottery if there is a tie on the second decimal place.

Figure C.4: Comparison of AI Relevance Distributions

(a) AI Producer Relevance

(b) AI User Relevance



Notes: The figures plot epanechnikov kernel density estimates with a bandwidth of 2. The blue is the distribution of AI relevance among all bachelor's graduates in the country. The red line is all priorities listed by all applicants in the CAS system 1996-2007. The green line is the restricted CAS dataset used for RD estimation.

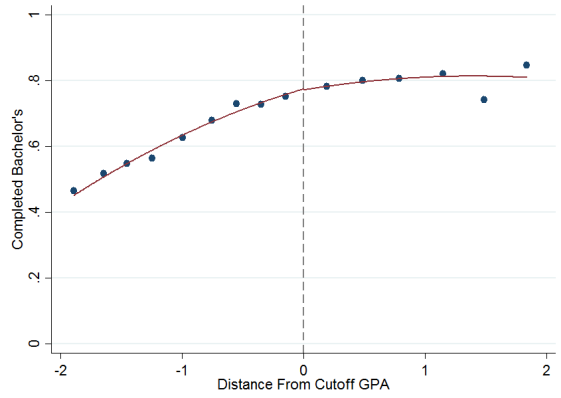
Table C.2: Correlation Between Earnings and AI Relevance in RD Samples

	High AI Preferred		Low AI Preferred	
	(1) Produce	(2) Use	(3) Produce	(4) Use
OLS Estimate (scaled)	1065.8	442.3	1547.5	708.4
	(75.43)	(34.63)	(106.9)	(115.3)
Observations	5990	8100	8633	6530

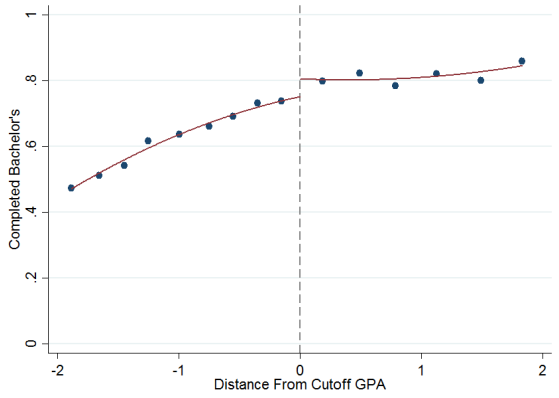
Notes: Estimation of a linear regression of earnings and AI relevance within RD samples. We use gender, age at application, and high school GPA as control variables. To compare with the results in Table 2c, we scale the OLS coefficient with the change in AI relevance of completed degrees from Table 2b. Subsample of applicants who eight years later has completed a bachelor's degree categorized with the AI relevance measure. Excluding applicants exactly on the cutoff.

Figure C.5: Completion of Degree and Threshold Crossing

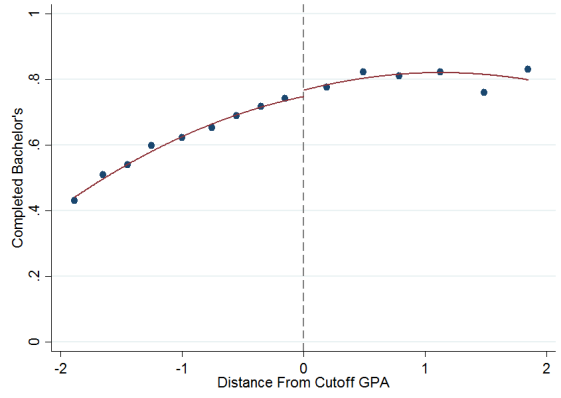
(a) Preferred Degree Is AI Producer Relevant



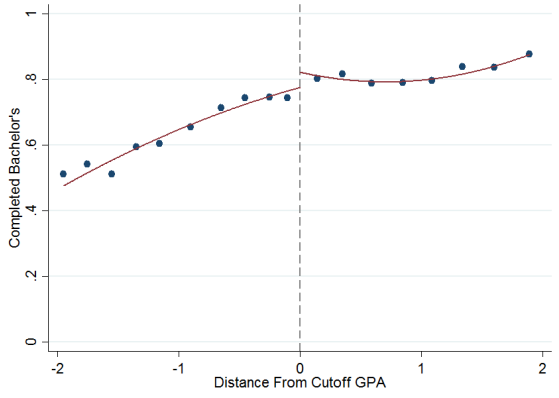
(b) Next-Best Is AI Producer Relevant



(c) Preferred Degree Is AI User Relevant



(d) Next-Best Is AI User Relevant



Notes: The figures plot completion rate of any bachelor's program, eight years after application, within bins of GPA minus cutoff, and fit estimated quadratics using all the underlying data. Excluding applicants exactly on the cutoff.

C.3 Causal Effects of Admitting Students to Majors With Higher Average Earnings

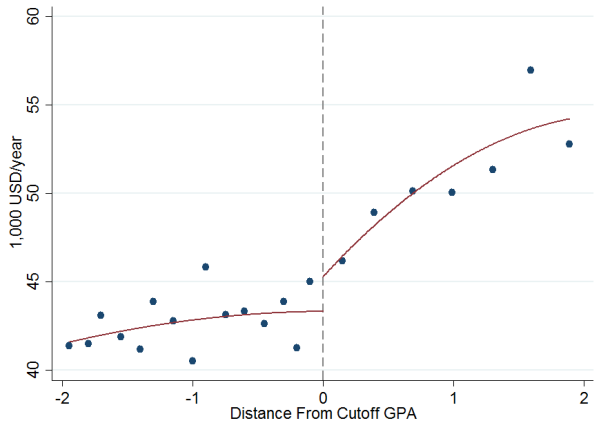
A useful benchmark for the causal earnings effects of AI relevance is to be more agnostic about what types of degrees that lead to higher earnings. In this section, we rerun our RD analysis, now ranking majors according to their average earnings.

Instead of forming pairs with different AI relevances, we form pairs where average earnings differ between the preferred and next-best priorities. When the preferred program has the highest expected earnings, the difference in earnings between the preferred and second-best programs is $15,562 on average. In comparison, when ranking on AI producer relevance, this difference is only $3,788. Put differently, the agnostic ranking pushes students between majors of very different average earnings.

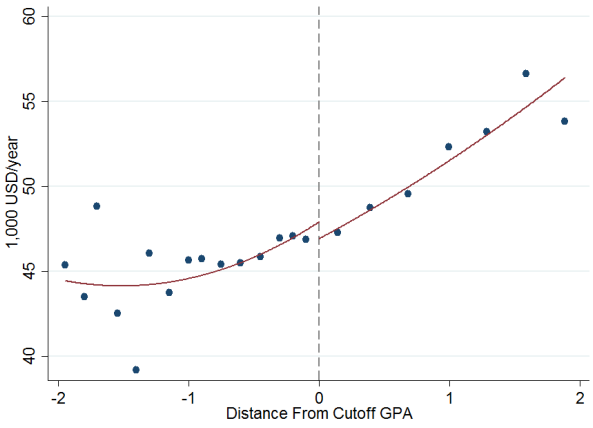
However, in Figure C.6 and Table C.3, we estimate that the causal effects of being admitted to a higher (lower) earnings degree only increase average earnings by $2,098 ($-452). The estimates, which are not statistically different from zero, are smaller than the AI producer relevance premiums estimated in Table 2c, despite the larger difference in average earnings between priorities. These results suggest that pushing students according to AI producer relevance, yields larger causal earnings returns than simply pushing them toward degrees with higher average salaries. One interpretation of the discrepancy between the causal earnings effects and the average earnings differences between majors is that earnings premiums on the AI producer margin are less driven by selection than the general earnings premiums between college majors.

Figure C.6: Earnings and Threshold Crossing - Expected Earnings Ranking

(a) Preferred is Higher Earning



(b) Next-Best is Higher Earning



Notes: The figures plot the mean of earnings within bins of GPA minus cutoff, and fit estimated quadratics using all the underlying data. Excluding applicants exactly on the cutoff and restricting to earnings above zero. Earnings are the median of wage income plus self-employed business income in the years seven to nine years after application. Income variables are deflated using the CPI with 2015 as the base year.

Table C.3: Earnings and Threshold Crossing - Expected Earnings Ranking

	High Earning Major Preferred	Low Earning Major Preferred
	(1)	(2)
RD Estimate	2097.7	-451.8
	(1343.0)	(1189.7)
Mean below cutoff	42898.6	45486.0
Observations	12208	15850

Notes: Estimation of equation (1) with gender and age as control variables. Excluding applicants exactly on the cutoff. Bandwidth is +/- 2.0 grade points.