

**UBS Center Working Paper Series**  
Working Paper No. 10, August 2014

ISSN 2296-2751 (Print)  
ISSN 2296-2778 (Online)

# Employment Polarization and the Role of the Apprenticeship System

Michelle Rendall  
Franziska J. Weiss



## About the UBS International Center of Economics in Society

The UBS International Center of Economics in Society, UBS Center in short, is an Associated Institute at the Department of Economics of the University of Zurich. It was established in 2012, enabled by a founding donation by UBS, which the bank made on the occasion of its 150th anniversary. In view of the very generous donation, the university named the UBS Center after its benefactor. The new center of excellence serves two main aims, the first of which is to finance world-class research in economics on all levels, to be conducted at the University's Department of Economics. It thereby supports the department's ambition to further strengthen its position as one of the top economics departments in Europe and to make Zurich one of the best places for research in economics in the world. The Center's other aim is to serve as a platform for dialogue between academia, business, politics, and the broader public, fostering continuous knowledge transfer in the process. Delivering on these aims will in turn also contribute to the quality of Zurich, and Switzerland more generally, as leading locations for education and business.

# Employment Polarization and the Role of the Apprenticeship System

**Michelle Rendall**

**Franziska J. Weiss**

University of Zurich

Department of Economics

## Abstract

This paper studies the effects of the apprenticeship system on innovation and labor market polarization. A stylized model with two key features is developed: (1) apprentices are more productive due to industry-specific training, but (2) from the firm's perspective, when training apprentices, technological innovation is costly since training becomes obsolete. Thus, apprentices correlate with slower adoption of skill-replacing technologies, but also less employment polarization. We test this hypothesis on German regions given local variation in apprenticeship systems until 1976. The results shows no employment polarization related to apprentices, but similar displacement of non-apprentices as in the US.

**Keywords:** Apprentices, educational system, employment polarization, technology adoption

**JEL classification:** E24, J24, O33, R23

## Acknowledgements

We would like to thank, among others, Marios Angeletos, Fabrizio Zilibotti as well as seminar and conference participants at the University of Zurich, for valuable comments. Financial support from the European Research Council (ERC Advanced Grant IPCDP-229883) is gratefully acknowledged.

We are solely responsible for errors and omissions.

## Contacts

University of Zurich, Department of Economics, Schoenberggasse 1, CH-8001 Zurich

Michelle Rendall, [michelle.rendall@econ.uzh.ch](mailto:michelle.rendall@econ.uzh.ch)

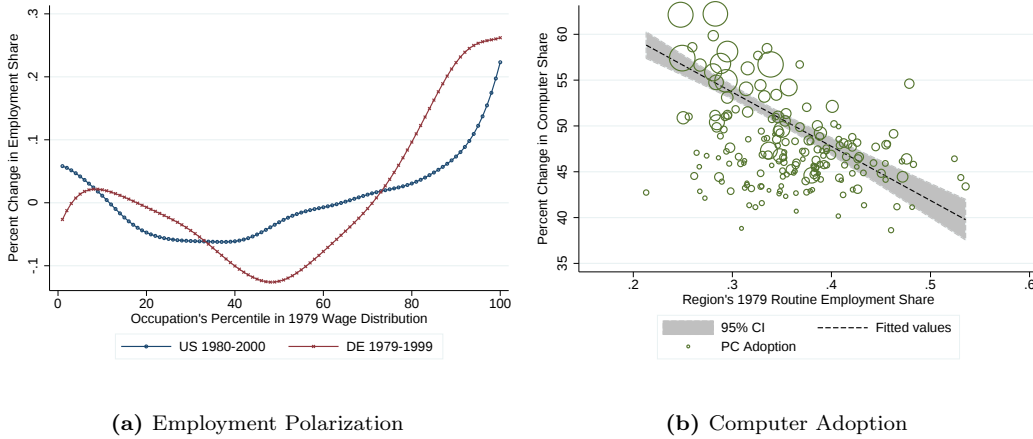
Franziska J. Weiss, [franziska.weiss@econ.uzh.ch](mailto:franziska.weiss@econ.uzh.ch)

## I. Introduction

One of the most significant changes in the labor market has been the increasing adoption of computer technology since the 1970s. Recent research has highlighted one specific impact of this ICT (information and communications technology) innovation: the displacement of middle income employment by capital. Autor et al. (2003) first documented the effect of computers, not only complementing the high-skilled, but also replacing middle-skilled jobs in the US. The authors decompose occupation requirements into three task types: manual (hand and finger dexterity), routine (repetitive) and abstract (analytical). Generally, the low, middle and high portions of the income distribution are linked to manual, routine and abstract tasks, respectively. Thus, computers most easily replace middle-skill tasks. Goos et al. (2011) document similar employment polarization across Europe and Senteleben and Wielandt (2012) document polarization in Germany. Figure 1a shows the change in employment shares for the US and Germany along the 1979 wage distribution of each country.

As polarization is present in both countries, it seems natural to conjecture that both countries have seen a rise in polarization driven by technical change in ICT. However, graphing computer adoption rates against the share of routine-intensive employment across German labor market regions (see Figure 1b) presents a puzzle in terms of the polarization hypothesis. Regions with the lowest routine employment share in 1979 have the most computer adoption. In contrast, in the US regions with the highest routine employment share have the largest per capita computer adoption rate (Autor and Dorn, 2013, see Table 3).

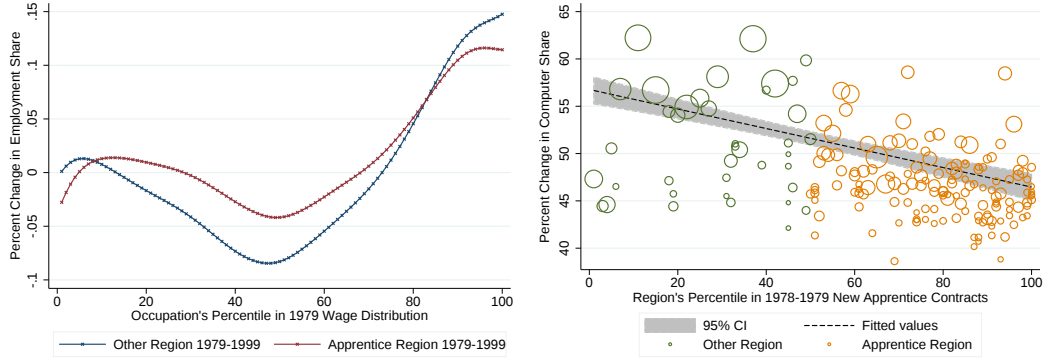
**Figure 1: Employment Polarization and Computers**



Note: US 1980/2000 census, German SIAB 1979/2000 samples and PC adoption from QCS 1999. For details see Appendix B.

In this paper we argue that differences in educational systems can resolve this puzzle. Figure 2b sorts regions along the percentile distribution in terms of new apprenticeship contracts in 1978/79 and graphs the share of personal computers used within each region by 1999.<sup>1</sup> The graph shows that regions with above average rates of new apprenticeship contracts in the 1978/1979 period (labeled “Apprentice

<sup>1</sup>New apprenticeship contracts are measured as the ratio of new contracts over the employed population within local labor markets.

**Figure 2: Employment Polarisation and Computers across Regions****(a) Employment by Apprentice-Intensity****(b) Computer Adoption by Apprentice-Intensity**

Note: German SIAB 1979/2000 samples, BiBB apprentice contracts 1978/1979, and QCS 1999 PC adoption. For details see Appendix B.

Regions”) have lower computer shares in 1999. The linear fit suggests a negative correlation of  $-0.45$ .

Figure 2a repeats Figure 1a for Germany by regions with above and below average new apprenticeship contracts from 1978/1979. Regions with above average ratios in apprenticeship contracts (or apprentice-intensive regions) have experienced significantly less employment polarization. Regions with little employment polarization are also regions with less computer adoption.

Autor and Dorn (2013) find that employment polarization has come with low-skilled service sector growth in the US. In comparison, structural transformation (or the rise of the service sector) has been considerably slower in Germany. In 1980 roughly 50 percent of Germany’s employment was manufacturing related (including agriculture), with this number dropping to roughly 38 percent in 2000. The US employment share in manufacturing fell from 34 to 25 percent during the same period. In terms of employment polarization, only the US distribution displays a marked U-shape with low wage occupations growing over time.

Goldin and Katz (2008) hint that slower output growth in Europe could be a function of the vocational education emphasis. Krueger and Kumar (2004) formalize this argument and suggest that since the 1980s Europe has lagged behind in terms of manufacturing productivity and in total output growth due to vocational educational systems. To our knowledge, we are the first to link the educational system with a notion of both skill biased technical change (SBTC) and task biased technical change (TBTC) and employment polarization (see Acemoglu and Autor, 2011, for a detailed survey on SBTC and TBTC).

The (dual) apprenticeship system incentivizes firms to invest in industry- or firm-specific skills of its workforce. Winkelmann (1997) suggests that the institutional structure that allows the apprenticeship system to function in Germany is likely not present in the US. That is, greater labor mobility, flexibility in wages and reduced layoff costs in the US make it riskier for firms to train workers in case of “poaching.” Acemoglu and Pischke (1998) show that the US is characterized by a high-turnover, low-training equilibrium. Thus, high school graduates (including associate degree holders or college dropouts) in the US are individuals that acquire general human capital at school, but receive very little specific training. In contrast, German ap-

prentices acquire specific skills. Skills are learned during the training period and not later, as apprentices have higher wages than unskilled workers in Germany, but have almost no returns to experience or tenure within the firm (for empirical evidence see Adda et al., 2006; Hanushek et al., 2011).

The apprenticeship costs, estimated to be about €15,000 per year per apprentice by the *Bundesinstitut fuer Berufsbildung* (BiBB), are mostly born by firms (Harhoff and Kane, 1993; Soskice, 1994). From the firm’s perspective, acquiring new machinery in areas where apprentices are employed is more costly because the prior training (specific skill) becomes obsolete and the comparative advantage of the apprenticeship is lost.

If ICT capital replaces non-college workers (as suggested by Autor et al. (2003)), firms in Germany, when compared to the US, have less incentives to adopt new technology and machines. Michaels et al. (2010) document the positive correlation between high-skill demand/ high-skilled wages and ICT adoption across countries. Van Ark et al. (2003) show that the EU is “lagging” the US in terms of ICT capital adoption. Although non-ICT adoption is very high in Europe, ICT has diffused slowly with no “catch up” observed.

In this paper, we develop a simple “skill-/task-replacing” model to demonstrate how non-college, routine task labor is prone to substitution by high-skilled workers using capital/machines. Building on the empirical fact that apprentice-skills are immobile, we extend the model to a spatial equilibrium setting where local labor markets have differential degrees of skill-specific workers (apprentices) similar to Autor and Dorn (2013). The model predicts that regions using apprentices (i.e. non-college labor with specific human capital, instead of general workers): (1) adopt fewer computers; (2) face slower employment polarization or displacement of routine work; and (3) realize a slower rate of structural transformation or a smaller rise in low-skilled services, since apprentices primarily work in the manufacturing sector (i.e. slower employment polarization). We evaluate these hypotheses using German social security data across local commuter zones.

Although Germany has an extensive apprenticeship system, there exists substantial variation in regional apprenticeship intensity. Before the ratification of the *Berufsbildungsgesetz* at the beginning of the 1970s and the introduction of the *Ausbildungsplatzfoerderungsgesetz* in 1976, the apprenticeship system was regionally fragmented. Only after the two oil shocks and the following deterioration of employment opportunities for young adults did the federal government make an effort to promote the apprentice-system nationwide, along with providing a systematic federal structure (or incentives) to educate the workforce.<sup>2</sup> Exploiting regional variation within a country, rather than using cross-country variation, avoids accounting for a myriad of other potentially important cultural and institutional differences.

The empirical results confirm the model’s hypotheses. That is, regions with high apprenticeship rates have less computer adoption, less employment polarization and a smaller rise in service employment. In line with Senfleben and Wielandt (2012), the results for other (non-apprentice) routine-labor are similar in magnitude to the US for the decades from 1980 to 2000 in terms of labor displacement. A region with 10 percentage points more routine-labor has a 2.19 percentage point larger displacement of routine-labor per decade. The results on computer adoption and service employment are in the same direction as in the US, but smaller in magnitudes.

<sup>2</sup>See Casey (1991), and see BiBB (1977); Soskice (1994) and references therein for a detailed discussion of the apprenticeship system history in Germany.

The remainder of the paper is organized as follows: Section II presents the theoretical model and derives the testable implications. Section III provides description of data used. Section IV presents the empirical results and compares them both across regions within Germany and in an international context. Lastly, Section V concludes.

## II. The Model

To model the interaction between apprenticeships and technological change, we develop a partial equilibrium model. We assume skill supplies are given, and the population is divided into non-college (apprentice or general education) and college graduates.<sup>3</sup> For simplicity, we refer to non-college workers as low-skilled and college graduates as high-skilled. These labels do not imply that the low-skilled acquire no skills, rather they acquire different skills, but the skill set is lower in terms of formal schooling. Firms choose how many machines to purchase and which type of worker to use for each task. Machines are produced by monopolists each period and technological progress is exogenous.

### FINAL GOOD PRODUCTION

The model is loosely based on Acemoglu and Autor (2011) and Acemoglu and Zilibotti (2001), with the addition of apprentices. That is, the model contains both, the concept of SBTC and TBTC. The unique final good is made up of a continuum of intermediate goods,  $[0, 1]$ ,

$$Y_t = \exp \left[ \int_0^1 \ln y_t(i) di \right].$$

We normalize the price of the final good to one each period. Each intermediate good can be produced using machines and labor,

$$y_t(i) = \max\{\alpha_\ell(i), \mathbb{1}_{(\ell=a|i \geq \underline{x})} \hat{\alpha}_\ell(i)\} \ell_t(i) + (\alpha_h(i) h_t(i))^\beta \left[ \int_0^{N_t} (\alpha_k(i))^\beta (k_t(i, v))^{1-\beta} dv \right],$$

where  $\alpha_j(i) > 0$  captures the skill-specific comparative advantage of labor type  $j$  in producing good  $i$ . The apprentices' productivity skills,  $\mathbb{1}_{(\ell=a|i \geq \underline{x})}$ , in contrast to "regular" low-skilled workers,  $\ell_t$ , are captured by  $\hat{\alpha}_\ell(i) > \alpha_\ell(i)$ . However, apprentices are more productive only for a subset of intermediate goods,  $i \geq \underline{x}$ . Very low-skill production (e.g. sweeping), do not benefit from apprenticeship training. The relative wage differential between apprentices and non-apprentices in the data suggests apprentices' comparative advantage to be largest in middle-wage production. The high-skilled productivity is constant across all goods,  $\alpha_h(i) = \alpha_h \forall i \in [0, 1]$ , but requires capital,  $k(i, v)$ , to be productive. Capital's productivity varies over the interval,  $\alpha_k(i)$ .

### DEFINITION OF GOODS PRODUCTION

Each intermediate good has manual-, routine- and abstract-task components. Goods are sorted on the interval  $i \in [0, 1]$  from mostly manual to abstract, with

<sup>3</sup>For simplicity we assume skill supplies to be fixed. Incorporating an endogenous education decision does not change the results of employment polarization qualitatively.

routine in the middle. The manual-intensive component is decreasing over the unit interval. The routine-intensive component is an inverted U-shape on the interval of intermediate goods. Thus, only goods in the middle have a large routine component that is easily replaceable by machines. The abstract component is increasing over the unit interval.

Production processes at the low-end,  $i \in [0, \underline{x})$ , require only low-skills, *LST*. The middle interval,  $i = [\underline{x}, \bar{x}]$ , with a large routine component, is abbreviated by *RT*. For this interval, apprenticeships provide additional productivity, although these tasks can be done by apprentices or high-skilled labor (using machines). At the upper range,  $i \geq \bar{x}$ , production is abstract and complex. The high-skilled are most productive. Assumption 1 formalizes this idea in terms of comparative advantages.

**Assumption 1.**

$\frac{\alpha_\ell(i)}{\alpha_h}$  and  $\frac{\hat{\alpha}_\ell(i)}{\alpha_h}$  are continuously differentiable and strictly decreasing in  $i$ . In addition, machine productivity is highest in the routine interval, i.e.  $\frac{\partial \alpha_k(i)}{\partial i} < 0 \forall i \in [0, 1]$ .

For simplicity of derivations below, we assume that  $\alpha_k(i)$  takes two distinct values, such that Assumption 1 is still satisfied. The productivity of capital is greater on the *RT* interval than above it. That is,  $\alpha_k(i) = \alpha_{k1} \equiv 1 \forall i \in [\underline{x}, \bar{x}]$  and  $\alpha_k(i) = \alpha_{k2} < 1 \forall i \in (\bar{x}, 1]$ .

Furthermore, it is assumed that the more abstract a production process is, the more difficult it is for low-skilled labor to be productive. Assumption 2 summarizes this concept.

**Assumption 2.**

$\frac{\partial^2 \alpha_\ell(i)}{\partial i^2}, \frac{\partial^2 \hat{\alpha}_\ell(i)}{\partial i^2} < 0 \forall i \in [0, 1]$ .

Given Assumption 1, there is perfect sorting between low- and high-skilled labor, with  $J_t$  as the sorting threshold. Since we are interested in studying the effect of technology displacing low-skilled labor from the routine interval, Assumption 3 guarantees that the economy's starting point is within the potential displacement region.

**Assumption 3.**

*The threshold on labor is within the RT regions for all time-periods,  $J_t \in [\underline{x}, \bar{x}] \forall t$ .*

## MACHINE PRODUCTION

The level of technology,  $N_t$ , is exogenously given (e.g. adopted from the world technological frontier). The market for machines is perfectly competitive such that the price equals the marginal cost. The production of one machine requires  $(1 - \beta)$  units of the intermediate good  $y_t(i)$ ,

$$\max_{k_t(i,v)} p_t^k(i,v) k_t(i,v) - (1 - \beta) p_t(i) k_t(i,v).$$

The price,  $p_t^k(i,v) = p_t^k(i) = (1 - \beta) p_t(i)$ , follows from the zero-profit condition.

### A. The Static Equilibrium

As in Acemoglu and Autor (2011) and Acemoglu and Zilibotti (2001), given Assumption 1, there is perfect sorting, with low-skilled labor working only in the interval  $[0, J_t)$ . Lemma 1 summarizes the production structure.

**Lemma 1.** *For any equilibrium, there is a threshold,  $J_t$ , in the interval  $\underline{x} < J_t < \bar{x}$ , such that for any  $i < J_t$ ,  $h_t(i)$ ,  $k_t(i, v) = 0$ , and for any  $i \geq J_t$ ,  $\ell_t(i) = 0$ .*

#### PRODUCTION

Producers of the final good are price takers and maximize profits, taking the price of their product ( $p_t(i)$ ), wages ( $w_{\ell t}$ ,  $w_{ht}$ ) and the price of machines ( $p_t^k(i, v)$ ) as given. Maximizing profits gives the demand for intermediate machines for each vintage type,  $v$ , and for each intermediate good,  $i$ ,

$$k_t(i, v) = \alpha_h \alpha_k(i) h_t(i).$$

The demand for machines at each production process  $i$  is,

$$(1) \quad k_t(i) = \int_0^{N_t} k_t(i, v) dv = \alpha_h \alpha_k(i) h_t(i) N_t.$$

Technological progress results in capital deepening. Given the productive schedule of capital, in Assumption 1, goods production in the routine interval compared to production above the routine interval uses relatively more capital (if the goods are produced by high-skilled labor). Given the demand for machines and Lemma 1, the production of good  $y_t(i)$  is,

$$\begin{aligned} y_t(i) &= \alpha_\ell(i) \ell_t(i) \quad \text{if } 0 \leq i < J_t \\ y_t(i) &= \alpha_h \alpha_k(i) h_t(i) N_t \quad \text{if } J_t \leq i \leq 1. \end{aligned}$$

For any equilibrium, the marginal product of each skill group must be equalized across all goods produced. This means price differences must exactly offset productivity differences,

$$(2) \quad p_t(i) \alpha_\ell(i) = p_t(i') \alpha_\ell(i') := P_{\ell t},$$

and

$$(3) \quad p_t(i) \alpha_h \alpha_k(i) = p_t(i') \alpha_h \alpha_k(i') := P_{ht}.$$

#### NO ARBITRAGE CONDITION ACROSS SKILLS

Due to the Cobb-Douglas production function, expenditures,  $p_t(i) x_t(i)$ , are equalized across all goods, which implies that the low- and high-skilled are equally distributed across goods,  $\ell_t(i) = \frac{L_t}{J_t}$  and  $h_t(i) = \frac{H_t}{1-J_t}$ . Hence, the following cost condition must hold in equilibrium,

$$\frac{P_{ht} H_t N_t}{1 - J_t} = \frac{P_{\ell t} L_t}{J_t}.$$

Substituting prices from equations (2) and (3) yields the “no arbitrage” condition, which pins down the equilibrium threshold,

$$(4) \quad \frac{\alpha_h \alpha_k(J_t) H_t N_t}{1 - J_t} = \frac{\alpha_\ell(J_t) L_t}{J_t}.$$

Although the model has implications for wages, the study of wages is beyond the

scope of this paper due to the country specific issues outlined in Section IV.D.

### B. Technical Progress

The world technological frontier advances at the exogenous rate  $(1 + \gamma) > 1$  or  $N_t = (1 + \gamma)^t N_0$  with  $N_0 = 1$ . The technical knowledge up to the frontier is assumed to be perfectly available to each producer, whereupon he/she decides on how much to invest in machines of each vintage.

### EMPLOYMENT DIVISION

Considering a one-time change in  $N$  we totally differentiate the no arbitrage condition in equation (4) with respect to  $N_t$ ,

$$(5) \quad \frac{dJ_t}{d \ln(N_t)} = \left( \frac{\partial \omega(J_t)}{\partial J_t} - \frac{1}{1 - J_t} - \frac{1}{J_t} \right)^{-1} < 0,$$

where  $\omega(i) \equiv \ln(\alpha_\ell(i)) - \ln(\alpha_h \alpha_k(i))$  is the productivity differences between the low- and high-skilled. By Assumption 1,  $\omega' < 0$ . Driven by skill-biased technical change, the range of goods produced by high-skilled workers increases at the expense of low-skilled workers. From the threshold cross-derivative, under Assumptions 1 and 2, the negative threshold growth accelerates as the threshold decreases,

$$(6) \quad \frac{d^2 J_t}{d \ln(N_t) d J_t} = - \left( \frac{dJ_t}{d \ln(N_t)} \right)^2 \left( \frac{\partial^2 \omega(J_t)}{\partial J_t^2} + \frac{1 - 2J_t}{(J_t(1 - J_t))^2} \right) > 0 \text{ for all } J_t > J^*.$$

Note, the threshold  $J_t$  has to be above a critical value,  $J^*$ , where  $J^* \in (0, \frac{1}{2})$  (see appendix A for the derivation of  $J^*$ ). Initially, it is very costly to spread high-skilled labor over more production processes. However, the marginal cost of spreading high-skilled labor over additional goods decreases with the threshold. I.e. spreading high-skilled labor from one to two goods is more costly than spreading it from two to three goods.

### MACHINE INVESTMENT

Total machine demand,  $X_t$ , is given by,

$$(7) \quad X_t = \int_{J_t}^1 \int_0^{N_t} k_t(i, v) dv di = \alpha_h N_t \frac{H}{1 - J_t} [(\bar{x} - J_t) + (1 - \bar{x}) \alpha_{k2}].$$

New technologies, a rise in  $N_t$ , increase machines' relative productivity making firms adopt more machines,

$$(8) \quad \frac{d \ln(X_t)}{d \ln(N_t)} = \frac{\partial \ln(X_t)}{\partial \ln(N_t)} + \frac{\partial \ln(X_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} > 0.$$

The first term is the direct productivity effect: each intermediate good uses more machines. The second term is the indirect effect: more goods are done by machines (and high-skilled labor). Since the direct effect is always positive, irrespective of the threshold, the following analysis only focuses on the second term. The effect of a

threshold change on machine adoption is,

$$(9) \quad \frac{\partial \ln(X_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} = - \frac{dJ_t}{d \ln(N_t)} \left( \frac{1}{\bar{x} - J_t} - \frac{1}{1 - J_t} \right) > 0.$$

There are two effects, (1) the interval in which machines are used increases and routine-labor is replaced by machines and high-skilled labor (i.e.  $J_t$  decreases); but (2) each machine is used by fewer high-skilled workers (the last term in Equation (9)). Given Assumption 1, machines are more productive in the  $RT$ -interval and the first effect always outweighs the second. That is, in the  $RT$ -interval machines can do most of the production with the “help” of a few high-skilled workers, while in the abstract interval machines need relatively more high-skilled workers to produce each unit of intermediate good.

#### ROUTINE-LABOR DISPLACEMENT

The labor share performing routine occupations is defined as  $L_{RT,t} := \frac{L_t}{J_t}(J_t - \underline{x})$ . The interval  $(J_t - \underline{x})$  is a subset of the total interval of  $RT$ . Strictly speaking, the model suggests that high-skilled workers and machines perform all tasks in the interval  $i \in [J_{h,t}, 1]$ , even though machines become increasingly productive at more routine tasks. Therefore, routine-labor should also include high-skilled labor by Assumption 3. However, each good  $i$  has three components: manual, routine and abstract. With only labor, labor must cover all three components simultaneously. For example, a secretary does manual, as well as routine and abstract tasks for the production of intermediate good  $i = j$ . However, if there is both labor and machines, the routine component from  $i \in [J_t, \bar{x}]$  might always be performed by machines and the non-routine component by high-skilled labor. E.g., the addition and subtraction of numbers are performed by machines and the more analytical tasks are performed by high-skilled labor using machines. Therefore, we abstract from high-skilled labor in the  $RT$  interval.<sup>4</sup> With technical progress, under Assumptions 1 and 3, the change in routine-labor is,

$$(10) \quad \frac{\partial \ln L_{RT,t}}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} = \left( \frac{\underline{x}}{(J_t - \underline{x})J_t} \right) \frac{dJ_t}{d \ln(N_t)} < 0.$$

Technological progress leads to a fall in routine employment. As some labor shifts to the production of low-skilled goods,  $i < \underline{x}$ , the relative fall in the number of routine labor is always larger than the increase in the number of workers per good  $i$ .

#### LOW-SKILLED SERVICE SECTOR EMPLOYMENT

If as in Autor and Dorn (2013), we label the lowest goods as services, then labor in low-skilled services,  $L_{LST,t} := \frac{L_t}{J_t}\underline{x}$ , is all labor in the interval of  $i = [0, \underline{x}]$ . The displacement of labor to production processes below  $\underline{x}$  implies that employment in low-skilled services increases upon ICT capital adoption by,

$$(11) \quad \frac{\partial \ln(L_{LST,t})}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} = - \frac{1}{J_t} \frac{dJ_t}{d \ln(N_t)} > 0.$$

<sup>4</sup>Alternatively, we could write a model combining the pure substitution effect of machines and middle-skilled tasks as in Acemoglu and Autor (2011), including the complementarity between machines and high-skilled tasks, as in Acemoglu and Zilibotti (2001). As this is a stylized model, we use a simpler approach by only modeling one part explicitly.

## THE EFFECTS OF TECHNICAL PROGRESS

Proposition 1 summarizes the effects of technological progress on machine adoption and the allocation of labor.

**Proposition 1.** *Upon arrival of new (ICT) technologies, a rise in  $N_t$  causes:*

- 1) *The adoption of more machines;*
- 2) *The displacement of labor in the routine interval; and*
- 3) *An increase of low-skilled service labor.*

## C. Spatial Equilibrium

To discuss differences across apprentice-intensive and general-education regions, we introduce an integrated spatial equilibrium model without labor mobility.<sup>5</sup> Following Autor and Dorn (2013), we assume that goods from each region are perfect substitutes, which ensures that goods' prices are equated in equilibrium.

The model analyzes two regions, with either apprentices or low-skilled workers. The only difference between regions is the relative productivity of the low-skill group,  $\hat{\alpha}_\ell(i) > \alpha_\ell(i)$ . We denote the “specific apprentice productivity” as  $\lambda$ , where  $\hat{\alpha}_\ell = f(\alpha_\ell, \lambda)$  and  $\frac{\partial \hat{\alpha}_\ell}{\partial \lambda} > 0$ . *Ceteris paribus*, a region with apprentices has a higher threshold  $J_t$ ,

$$(12) \quad \frac{dJ_t}{d\lambda} = -\frac{1}{\hat{\alpha}_\ell(J_t)} \frac{\partial \hat{\alpha}_\ell(J_t)}{\partial \lambda} \frac{dJ_t}{d \ln(N_t)} > 0.$$

Apprentice productivity,  $\hat{\alpha}_\ell(i)$ , can take any functional form as long as (1) Assumptions 1 and 2 are satisfied, and (2) regions with apprentices have a slower fall in the threshold upon technology adoption,

$$(13) \quad \frac{d \frac{dJ_t}{d \ln(N_t)}}{d\lambda} > 0,$$

or

$$(14) \quad \left( \frac{d^2 \omega(J_t)}{dJ_t^2} + \frac{1 - 2J}{(J_t(1 - J_t))^2} \right) \frac{1}{\hat{\alpha}_\ell(J_t)} \frac{d\hat{\alpha}_\ell(J_t)}{d\lambda} \frac{dJ_t}{d \ln(N_t)} > \frac{d^2 \omega(J_t)}{dJ_t d\lambda}.$$

PROOF:

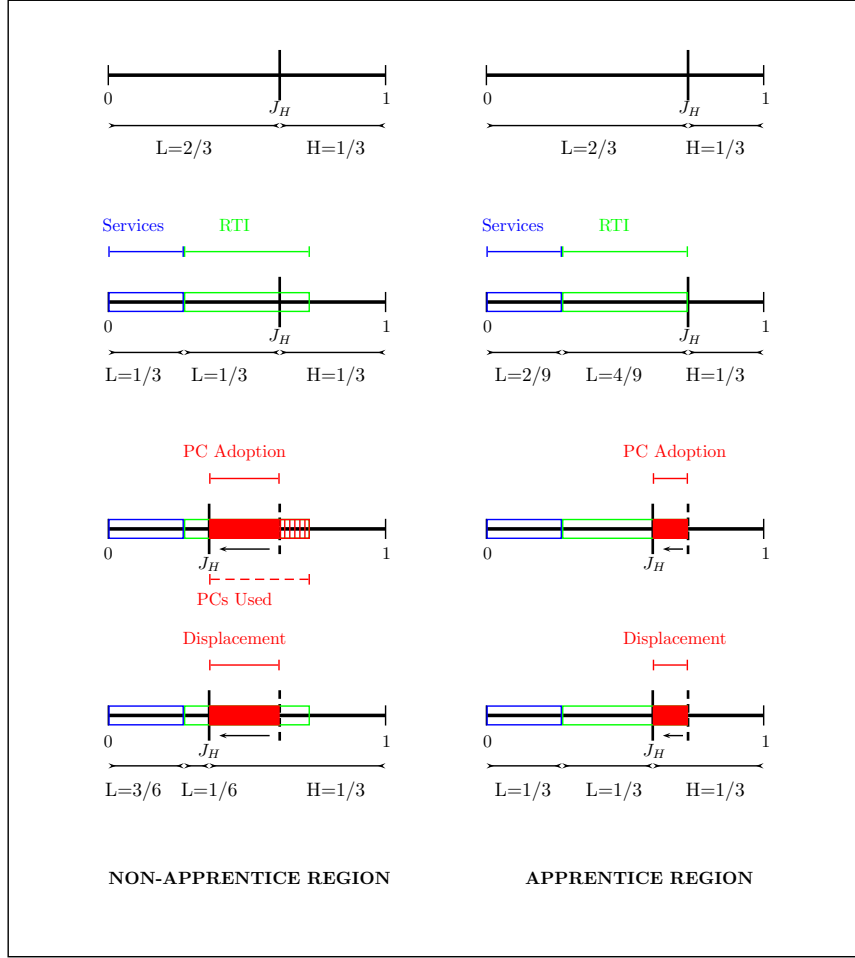
See Appendix A.

The left hand side is strictly positive as long as  $J_t > J^*$  (see Equation (6)). Hence the condition holds whenever  $\frac{d^2 \omega(J_t)}{dJ_t d\lambda} < 0$ . One productivity schedule that fulfills this condition is  $\hat{\alpha}_\ell(i) = \alpha_\ell(i) \cdot \lambda(i)$  with  $\frac{\partial \lambda(i)}{\partial i} < 0$ . The qualitative results are independent of the exact  $\hat{\alpha}_\ell(i)$ -schedule, as long as Equation (14) holds.

The cross derivatives of Equations (9), (10) and (11) with respect to the apprentice productivity,  $\lambda$ , summarize the technology effect between apprentice- and other regions.

<sup>5</sup>Dustmann and Pereira (2008) show that job mobility (wage growth and returns to experience) is substantially lower in Germany compared to the UK. Adda et al. (2006) further document that apprentices, in particular, are less mobile and overall job mobility is substantially lower in Germany compared to the US. Since job mobility is a prerequisite for regional mobility, these facts also suggest low overall mobility.

Figure 3: Hypotheses 1 and 2



Note: The figure displays the change in the occupational structure across regions facing an increase in ICT capital. This effect is displayed for a non-apprentice region in the left panel and for an apprentice-intensive region in the right panel.

### TECHNOLOGY EFFECTS ACROSS REGIONS

The differential effect of technology on the dependent variable,  $Z_t \in \{X_t; L_{RT,t}; L_{LST,t}\}$ , from Equations (9)-(11) is,

$$(15) \quad \frac{d\left(\frac{\partial \ln(Z_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)}\right)}{d\lambda} = \frac{\partial \left(\frac{\partial \ln(Z_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)}\right)}{\partial J_t} \frac{dJ_t}{d\lambda} + \frac{\partial \left(\frac{\partial \ln(Z_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)}\right)}{\partial \lambda}.$$

Using the results from Equations (5), (6), (12), Condition (14) and Assumptions 1 - 3, it can be shown that apprentice-regions have less machine adoption,  $X_t$ , less routine-labor displacement,  $L_{RT,t}$ , and less growth in low-service employment,  $L_{LST,t}$  given a threshold in the  $RT$ -region,  $J_t > J^*$ . PROOF:

See Appendix A.

Figure 3 visualizes the effect of machine adoption and routine-labor displacement. The differential speed of machine adoption across regions is only driven by differences in the substitution channel (of the low-skilled). The complementarity effect, condi-

tional on regions having the same high-skilled labor share, is identical in apprentice- and other regions. Intuitively, for the results on labor, apprentices are more productive than other low-skilled workers and, therefore, the opportunity cost of replacing them is larger. Consequently, there is less labor available to grow the low-skilled service sector. Note that the growth in services is the inverse of labor displacement.

#### D. Testable Implications and Empirical Specification

Using the results from Equation (15), we derive three testable implications across commuting zones for the dependent variable,  $Z_t \in \{X_t; L_{RT,t}; L_{LST,t}\}$ .

##### EMPIRICAL SPECIFICATION

The data does not provide a direct measure of technological change,  $N_t$ . In the model, the change in the threshold parameter ( $J_t$ ) only changes with technological progress. Therefore,

$$\frac{dJ_t}{d\ln(N_t)} \equiv \frac{dJ_t}{dt}.$$

Then the effects of technology on the dependent variable ( $Z_t$ ) is,

$$\frac{\partial \ln(Z_t)}{\partial J_t} \frac{dJ_t}{d\ln(N_t)} \approx \frac{\Delta \ln(Z_t)}{\Delta J_t} \frac{\Delta J_t}{\Delta t}.$$

Ignoring apprentices, the change in  $J_t$  over time only depends on the region's initial  $J_0$ . Conditional on  $J_0$ , regions have the same change in the threshold value and, hence, in the dependent variable,

$$\frac{\Delta \ln(Z_t)}{\Delta J_t} \frac{\Delta J_t}{\Delta t} \equiv \Delta(\ln(Z_t) - \ln(Z_0)) \Big|_{J_0}.$$

Thresholds,  $J_t$  are an abstract concept and not observed in the data. However, a region's routine-labor share is observable,  $L_{RT,0}$ . Without apprentices, the routine-labor share is determined by the threshold - the higher the threshold, the higher the routine-labor share. The threshold can be replaced by the routine-labor share,  $L_{RT,0}$ , yielding,

$$\Delta(\ln(Z_t) - \ln(Z_0)) \Big|_{J_0} \approx \Delta(\ln(Z_t) - \ln(Z_0)) \Big|_{L_{RT,0}}.$$

Since all regions face the same global technology frontier, the initial routine-labor share,  $L_{RT,0}$ , captures all the variation in the threshold  $J_0$  (given the two types of labor, low- and high-skilled must equal one). Therefore, the baseline regressions corresponding to Equations (9)-(11) are,

$$(16) \quad \Delta \ln(Z_{t,j}) = \beta_0 + \beta_1 L_{RT,0,j} + \epsilon_j.$$

When differentiating between apprentices and other low-skilled labor, the routine labor share,  $L_{RT,0}$  does not uniquely define the threshold,  $J_0$ . A region's threshold is determined by (1) the share of high-skilled labor and (2) the composition of low-skilled labor (apprentices or others). Only conditional on the high-skilled labor share does a region with more apprentices have a higher threshold given technology.

Equation (15), can be rewritten as,

$$\frac{d\left(\frac{\partial \Delta \ln(Z_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)}\right)}{d\lambda} = \frac{d\Delta(\ln(Z_t) - \ln(Z_0))}{d\lambda} \Big|_{L_{RT,0}, H_0}.$$

The regressions show the differential effect of apprentice and other low-skilled labor across regions are,

$$(17) \quad \Delta \ln(Z_{t,j}) = \beta_0 + \beta_1 L_{RT,0,j}^A + \gamma_1 L_{RT,0,j}^O + H_{0,j} + \epsilon_j,$$

where “A” stands for apprentice labor and “O” for other low-skilled labor, with  $H_{0,j}$  being the region’s initial high-skilled share.<sup>6</sup>

#### TESTABLE IMPLICATIONS

The resulting testable implications of Equation (15) are:

**Hypothesis 1.** *Conditional on the high-skilled labor share, regions with an apprentice-intensive industry structure adopted fewer computers ( $X_t$ ) over time.*

**Hypothesis 2.** *Apprentice-intensive regions, conditional on the high-skilled labor share, have less displacement of routine-labor ( $L_{RT,t}$ ) as new ICT technology is invented.*

**Hypothesis 3.** *Apprentice-rich regions experience a smaller rise in low-skilled services ( $L_{LST,t}$ ) over time, conditional on the high-skilled labor share.*

### III. German Regional Data

This section summarizes our construction of key variables and data sources. Further detail is provided in Appendix B.

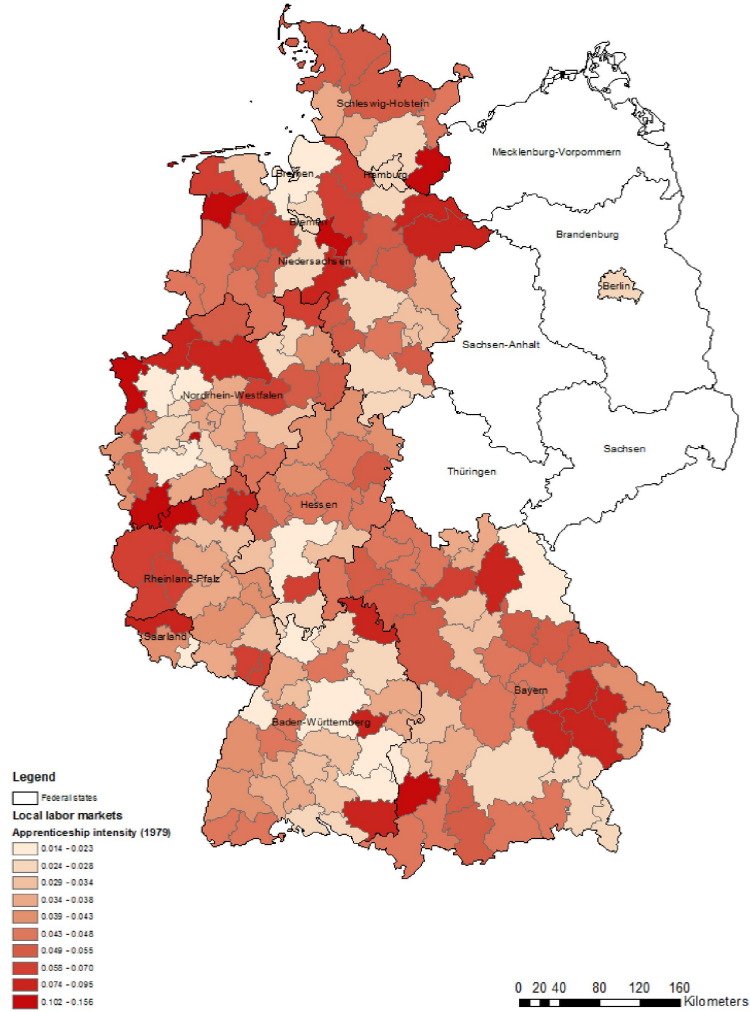
#### A. Data Sources

Two main data sources are used in this paper: (1) the Sample of Integrated Labor Market Biographies - Regional File 1975-2008 (SIAB); and (2) the BIBB/IAB Qualification and Career Survey 1979 and 1999 (QCS). The SIAB sample provides detailed individual-level characteristics, such as education, region of work, nationality, and labor market experience (e.g. occupational status and wages) (see Dorner et al., 2011, for details). The sample used consists of all workers subject to social security payments, aged 17 to 62. Following Dustmann et al. (2009), all workers are weighted by part-time or full-time work given limited information on hours worked in the regional sample.

The QCS is used to construct occupation-specific computer usage. We use task measures computed by Autor and Dorn (2013) for the US to compute task-intensity across the German working population. We do so in order to make the research comparable to US studies, even though the QCS provides information on tasks performed by each individual. Although the German data identifies routine, manual, and abstract tasks, the questions are not identical to the US survey and, therefore,

<sup>6</sup>Unlike the theory, regions do not perfectly sort into apprentice versus non-apprentice regions, as regions have different mixes of the two types of educational systems. Therefore, it is necessary to differentiate between apprentice and other low-skilled labor.

Figure 4: Germany Apprenticeship Intensity in 1979



Note: BiBB apprentice contracts 1978/1979.

will not necessarily capture the same information. Nonetheless, the German task measure provides very similar results. A comparison between the German and US task measures, applied to the SIAB sample, is provided below.

To compute regional variation we use the official definition of local labor markets (or commuting zones) from the “Bundesinstitut fuer Bau-, Stadt- und Raumforschung,” which is quasi identical to the classification of the “Gemeinschaftsaufgabe Verbesserung der regionalen Wirtschaftsstruktur” (for details see Koller and Schwengler, 2000). These zones are based on economic geography, accounting for commuter flows and commuting time (see Eckey, 1988; Eckey and Klemmer, 1991). Using the SIAB dataset, for West Germany only, results in 182 commuting zones. Figure 4, sorting local labor markets into deciles of new apprenticeship contracts in 1978/1979, shows the initial regional variation across West-Germany. There is no clear North-South, or West-East pattern in the late 1970s.

Thus, in our empirical analysis, we use initial regional variation in apprentice

shares and routine employment shares to determine the interaction between education, ICT adoption and labor market polarization. As a robustness check we later also control for initial industry structure (see Appendix C). This robustness check underscores that the results presented below are driven by apprentice variations and not that apprenticeship variations are a result of industry structure and industry structure has driven computer adoption.

### B. Apprentices

Apprenticeship rates in Germany are high and stable over time. The vocational education reports from 1977 and 2011 (BiBB, 1977, 2011) document 496,000 new apprenticeship contracts in 1976 (67 percent of all secondary graduates) and 468,410 new contracts in West Germany in 2010 (65 percent of secondary graduates). In West Germany, 95 percent of apprenticeships in 2010 were financed by firms and the remainder by public funds.

While initial apprenticeship numbers are large, a considerable fraction eventually switch industries and occupations, making most of the specific knowledge obsolete. BiBB (1977) reports that about 40 percent of male apprentices between 1955 and 1970 switched their broad sector of work. About half of them document that their specific skills became obsolete (see also Adda et al., 2006). Using the QCS sample, we find about 50 percent of apprentices switching industry and about 31 percent switching the broad sector (services to non-services). In the empirical analysis, we label workers as apprentices if they, at some point, completed an apprenticeship in the same broad sector they are currently employed in. This is consistent with the model's assumption that apprentices only have additional productivity in the *RT* sector.

### C. Measuring Inputs

Since the data is available on an individual level we measure inputs to the production process on an occupational level rather than the intermediate goods level as in the model. However, for the analysis only the aggregate regional input levels matter and not how they are distributed across individual occupations. We make use of both the individual tasks from Autor et al. (2003) as well as the compounded routine measure computed as,

$$R_j = \frac{routine_j}{routine_j + manual_j + abstract_j},$$

for occupation  $j$ . Routine tasks are an average of routine-cognitive and routine-manual tasks. Abstract tasks are the average of non-routine personal and non-routine analytical. Manual tasks are the non-routine manual tasks (for further details see Autor et al., 2003). As in Autor and Dorn (2013), an occupation is labeled routine-task-intensive,  $RTI_j$ , if the occupation falls within the top one-third of the employment distribution in terms of the  $R_j$  measure,

$$(18) \quad RTI_j = \mathbf{1} [R_j \geq R^{P66}].$$

For the analysis of machine adoption we also compute the 66th percentile index for the separate task measures (manual, routine and abstract).

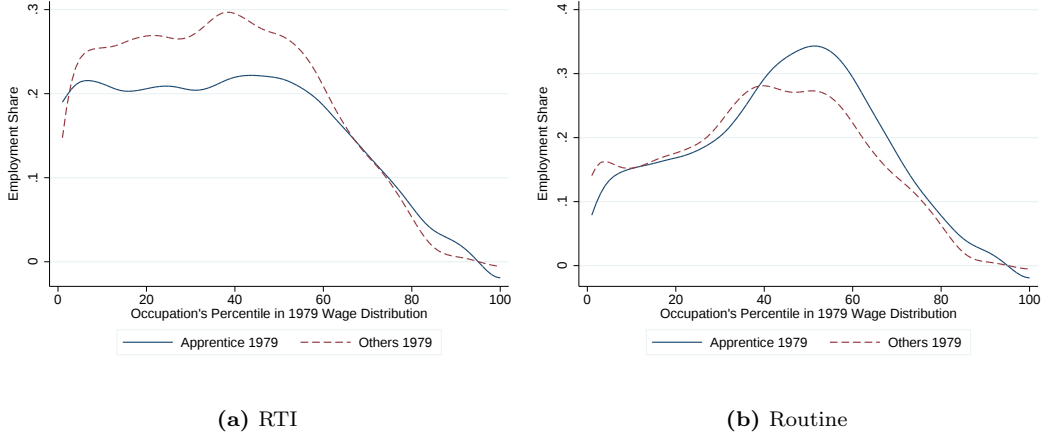
**Figure 5: Share of RTI Occupations (DOT vs BiBB) in 1979**

Note: German SIAB 1979/2000 samples, US DOT 1977 and QCS 1979. For details see Appendix B.

Using US task measures on German occupations might have its limitations. However, we are able to compare the measure to similar German measures. The QCS, similarly to the dictionary of occupational title (DOT) in the US, provides information on types of tasks done by workers. While the questions asked are not identical between the two surveys, the two measures provide very similar aggregate results. Figure 5 shows the share of employment within RTI-occupations for both the DOT and the QCS measures across the 1980's wage distribution. Most of the employment in RTI-occupations occurs in the middle of the wage distribution (for the US equivalent, see Figure 4 in Autor and Dorn, 2013). Given these similarities the results are provided using the US data, allowing for a direct comparison between this study and US studies.

It is important to establish that both apprentices and non-apprentices could potentially be replaced by ICT technology. If apprentices were to perform very different tasks (i.e. tasks that are not routine in nature), it would not be surprising that apprentices are not displaced by ICT. Figure 2 would simply be a product of apprentices being irreplaceable and not that acquired skills increase their relative productivity. Figure 6 graphs the share of routine employment from Figure 5 separately, by apprentices versus other workers. The left panel (Figure 6a) graphs the aggregate RTI measure in employment shares and the right panel (Figure 6b) graphs the 66th percentile employment share of occupation that are routine (individual task measure). Although apprentices perform less RTI tasks along the lower part of the wage distribution, the overall employment in RTI tasks is similar. Comparing only the routine component of RTI suggests that apprentices and other workers engage in virtually the same amount of routine employment, especially in the middle of the wage distribution where the majority of employment polarization has taken place. The difference in the aggregate RTI measures is driven by some apprentice occupations performing more manual tasks than other workers.

Comparing the distributions from Figures 1, 2 and 5, 6, German employment po-

**Figure 6: Share of Routine Occupations (Apprentices vs Others) in 1979**

Note: German SIAB 1979/2000 samples and US DOT 1977. For details see Appendix B.

larization is likely due to the displacement of routine occupations by ICT technology. However, given Figure 2, this polarization is slower in apprentice-intensive regions even though apprentices work similarly routine-intensive occupations as other low-skilled workers.

#### IV. Empirical Results

We separately test our three hypotheses on: (1) computer adoption, (2) routine-labor displacement, and (3) and service sector growth from Section II.D.

To do this formally, cross-section OLS regressions are estimated. The cross-sectional variation originates from the regional differences at the commuting zone level and the fragmented apprenticeship system prior to the mid-1970s. All variables are constructed on a regional labor market level. For example, the regional measure of routine-task-intensity,  $RTI$ , is defined as the share of employment within region  $j$  that works in RTI-occupations  $k$ ,

$$RTI_{jt} = \left( \sum_{k=1}^K L_{jkt} \cdot RTI_k \right) \left( \sum_{k=1}^K L_{jkt} \right)^{-1}.$$

##### A. Computer Adoption and Skills

To determine the validity of Hypothesis 1, we test whether the adoption of personal computers can be explained by the share of RTI-occupations and/or the share of abstract-occupations in a region. The former suggests ICT is substitutable with the low-skilled (or routine tasks) and the latter suggests ICT is complementary with the high-skilled (or abstract tasks).

The computer measure is tabulated from the 1999 QCS. We adopt the standard procedure of assuming that the share of computers in 1999 also captures the growth in computer utilization since 1979. I.e. computers were virtually absent in 1979.<sup>7</sup>

<sup>7</sup>Although the 1979 QCS does provide computer utilization information, we do not rely on this data since the survey design changed substantially since 1979. Defining computer adoption in 1999 as the change between computer utilization in 1999-1979 provides similar results.

The QCS allows us to compute computer adoption by occupations, and the SIAB provides the regional variation in occupational structure. That is, it is assumed that a given occupation adopts computers in a similar way, regardless of location. E.g., it does not matter whether the adoption is located in the South or the North.<sup>8</sup> For additional detail on the construction of the computer measure, see Appendix B.B2.

Without differentiating between apprentices and others, we regress the computer (PC) measure in 1999,  $PC_{j,99}$ , on the regional measure of routine-task-intensity,  $RTI_{j,79}$ ,

$$PC_{j,99} = \beta_0 + \beta_1 RTI_{j,79} + \psi_f + \epsilon_j.$$

This regression is derived from equation (9) in Section II and the coefficient on  $RTI_{j,79}$ ,  $\beta_1 > 0$ , should be positive. Each regression has controls for federal state fixed effects,  $\psi_f$ . Standard errors are clustered on the federal state level and all regressions are weighted by each periods' initial employment shares.

Alternatively, we also show the correlation decomposed by separate tasks,

$$PC_{j,99} = \beta_0 + \beta_1 Routine_{j,79} + \beta_2 Abstract_{j,79} + \beta_3 Manual_{j,79} + \psi_f + \epsilon_j.$$

Note, by the definition of task shares, the share of manual-, routine-, and abstract employment within a region do not necessarily add up to one. Occupations are only classified as either manual-, routine-, or abstract- if they meet the 66th percentile employment share threshold from Equation (18). In this specification, the coefficient on routine-labor should still be positive,  $\beta_1 > 0$ . In addition, the coefficient on abstract-labor should also be positive,  $\beta_2 > 0$  and if the complementarity effect is stronger than the substitution effect, the latter coefficient should be greater,  $\beta_2 > \beta_1$ . Since manual-labor is not easily replaced by computers, and also has little complementarity with computers, we expect the coefficient to be zero,  $\beta_3 = 0$ .

Table 1 shows the correlation between PC Adoption in 1999 and initial employment shares in routine-intensive jobs in 1979 across regions. Column (1) uses the compounded RTI measure, column (2) uses the individual routine measure, and column (3) uses the three separate measures to distinguish between substitution and complementarity effects. Column (1) and (2) suggest, counter to the theory, that more RTI- or routine employment is associated with less computer adoption. However, the results of column (3), when including all task measures separately, suggest that both the substitution-of-routine-skills and complementary-to-abstract-skills are present in Germany. The substitution effect is considerably smaller. A region with 10 percentage points more routine employment in 1979 has 1.4 percentage points more computer adoption, whereas 10 percentage points more abstract employment in 1979 results in 11.3 percentage points more computers in 1999. There is no effect of the manual component. Moreover, results in column (3) are in line with results reported by Senftleben and Wielandt (2012) even though this paper uses different task measures.

Having established the same overall effects in Germany as in the US, Table 2 explores the differential effect between apprentices and other workers. In doing so, we differentiate the share of workers in RTI-occupations into apprentice-intensive,

<sup>8</sup>Unfortunately, regional variation in the QCS is sparse, and does not allow for a robust measure on the regional level.

Table 1: PC Adoption and Tasks

| VARIABLES      | (1)<br>PC_99         | (2)<br>PC_99          | (3)<br>PC_99         |
|----------------|----------------------|-----------------------|----------------------|
| RTI            | -0.660***<br>(0.103) |                       |                      |
| Routine        |                      | -0.663***<br>(0.0682) | 0.143**<br>(0.0430)  |
| Abstract       |                      |                       | 1.133***<br>(0.0377) |
| Manual         |                      |                       | 0.0176<br>(0.0619)   |
| N              | 182                  | 182                   | 182                  |
| R <sup>2</sup> | 0.530                | 0.632                 | 0.818                |

**Notes:** Standard errors are clustered at the federal state level and displayed in parentheses. \*\*\* significant at 1 percent, \*\* significant at 5 percent, \* significant at 10 percent. All regressions include federal state fixed effects (11 federal states). All models are weighted by each periods' initial employment shares. PC adoption is measured in 1999, the independent variables in 1979.

*ARTI* and other occupations, *ORTI*,

$$\begin{aligned}
ARTI_{j,t} &= \left( \sum_{k=1}^K L_{j,kt} RTI_k \times APP_k \right) \left( \sum_{k=1}^K L_{j,kt} \right)^{-1}, \\
ORTI_{j,t} &= \left( \sum_{k=1}^K L_{j,kt} RTI_k \times (1 - APP_k) \right) \left( \sum_{k=1}^K L_{j,kt} \right)^{-1}, \\
&\Rightarrow ARTI_{j,t} + ORTI_{j,t} = RTI_{j,t},
\end{aligned}$$

where  $APP_k$  indicates if the job is apprentice-intensive and, thus, has an apprentice-share above the 66th percentile.

We see a larger complementarity than substitution effect. Therefore, the compounded *RTI*-measure is unable to pick up the positive effect of routine employment on computer adoption. Thus, the next results only focus on the decomposed task measures. The computer (PC) measure in 1999,  $PC_{j,99}$ , is regressed on the  $Routine_{j,79}$ , the  $Abstract_{j,79}$ , and the  $Manual_{j,79}$  employment shares, and a control for the high-skilled labor share,  $H_{j,t}$ :

$$\begin{aligned}
PC_{j,99} &= \beta_0 + \beta_1 ARoutine_{j,79} + \beta_2 AAbstract_{j,79} + \\
&\quad \gamma_1 ORoutine_{j,79} + \gamma_2 OAbstract_{j,79} + \delta Manual_{j,79} + \psi_f + \epsilon_j.
\end{aligned}$$

Even though equation (15) is conditional on the high-skilled share, the control is omitted here, as the correlation between “Abstract” and the high-skilled share is 0.926. Equation (15) suggests the coefficient on routine-labor should be smaller in the apprentice region,  $0 < \beta_1 < \gamma_1$ . Given equation (9), it is unclear what the relationship of coefficients on abstract labor,  $\beta_2$  and  $\gamma_2$  should be. The model, however, suggests that these effects would be similar across regions since high-skilled labor is equally productive everywhere.

The results in Table 2 are provided both by differentiating all task measures by apprentices and other worker (columns (1)), and only focusing on the routine differ-

ence (column (2)). In column (1) the coefficient on abstract employment is smaller

**Table 2: PC Adoption, Apprentices and Tasks**

| VARIABLES      | (1)<br>PC_99        | (2)<br>PC_99        |
|----------------|---------------------|---------------------|
| ARoutine       | 0.0044<br>(0.234)   | 0.016<br>(0.236)    |
| ORoutine       | 0.153**<br>(0.064)  | 0.154**<br>(0.0537) |
| AAbstract      | 0.858***<br>(0.160) |                     |
| OAbstract      | 1.585***<br>(0.268) |                     |
| Abstract       |                     | 1.150***<br>(0.056) |
| Manual         | 0.023<br>(0.067)    | 0.046<br>(0.098)    |
| N              | 182                 | 182                 |
| R <sup>2</sup> | 0.835               | 0.819               |

**Notes:** Standard errors are clustered at the federal state level and displayed in parentheses. \*\*\* significant at 1 percent, \*\* significant at 5 percent, \* significant at 10 percent. All regressions include federal state fixed effects (11 federal states). All models are weighted by each periods' initial employment shares. PC adoption is measured in 1999, the independent variables in 1979.

for apprentices than other workers, but the two measures are qualitatively comparable. Moreover, we cannot reject the null hypothesis that they are statistically equal at the 5 percent confidence level ( $F(1, 8) = 3.49$ ). The coefficient on manual employment is insignificant (they are still insignificant if estimated separately). In addition, for both column (1) and (2) when apprentices work in routine tasks, there is no correlation with computer adoption. Regions with 10 percentage points more other routine-labor adopt 1.5 percentage points more PCs. The full substitution effect from Table 1 is driven by other workers, not apprentices. This is precisely the prediction from the model.

Until now we have only shown a correlation between routine-labor and computers, but have not established if these workers are replaced by computers. The next section shows that these workers are in fact displaced.

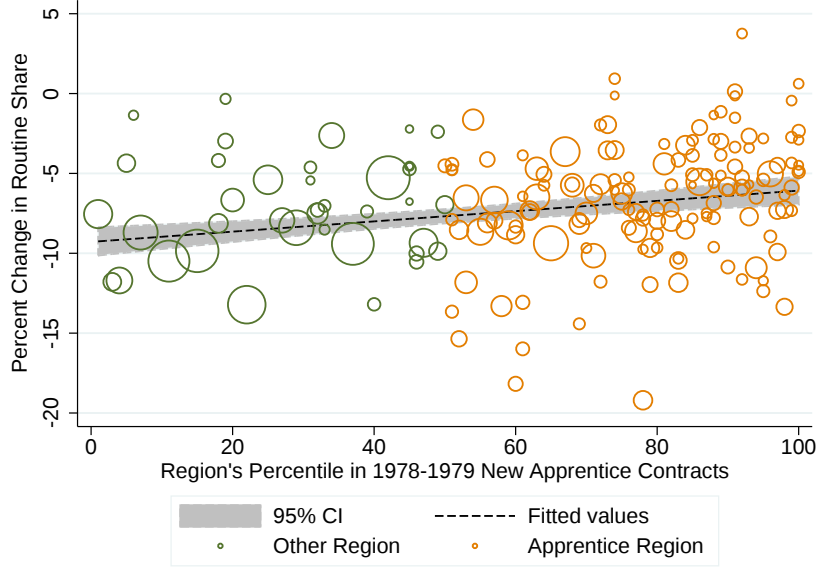
### *B. Routine Labor Displacement*

The model suggests that with the invention of ICT technologies, apprentice-intensive regions should experience less displacement of routine-labor. Visually, Figure 7 shows the correlation of apprentice-intensity and routine displacement in a similar manner as Figure 2b for computers. The negative raw correlation between the share of new apprentice contracts in 1978/1979 and the displacement of RTI-labor shares for the 1979-2008 time period is  $-0.26$ .

Ignoring apprenticeship differences, the regression of the change in routine task intensity (the employment share in routine-intensive occupations) on the period's initial routine task intensity, is

$$(19) \quad \Delta^{(t+10)-t} RTI_j = \beta_0 + \beta_1 RTI_{j,t} + \psi_f + \theta_t + \epsilon_j,$$

Figure 7: Displacement by Region and 1979 Apprentice Intensity



Note: German SIAB 1979/2000 samples and BiBB apprentice contracts 1978/1979.

which follows from equation (10). Given the theoretical framework and the evidence from the US, we expect the coefficient on  $RTI$ -labor to be negative,  $\beta_1 < 0$ .

All regressions have controls for federal state fixed effects and time fixed effects,  $\theta_t$ . Standard errors are clustered on the federal state level and results are weighted by the period's initial employment shares. For brevity, we report results for the compounded measure only, although results for the separate routine measure are comparable.<sup>9</sup> Table 3 column (1) is a stacked multi-period model, where the dependent variable is the decade change in the routine share. The coefficient in column (1) is negative but insignificant. Column (2) repeats column (1), but differentiating by apprentice and other employment. The change in employment is regressed on initial apprentice or other  $RTI$ -employment and the initial high-skilled labor share,  $H_{t,j}$ ,

$$(20) \quad \Delta^{(t+10)-t} RTI_j = \beta_0 + \beta_1 ARTI_{j,t} + \gamma_1 ORTI_{j,t} + \delta H_{t,j} + \psi_f + \theta_t + \epsilon_j.$$

The control for the high-skilled labor share follows directly from equation (15). Given the cross-derivative, the coefficient on routine-labor should be negative and, in absolute value, larger for other employment, i.e.  $\gamma_1 < \beta_1 < 0$ . Column (2) corroborates the theory, finding displacement for other workers, but not apprentices. Per decade, 10 percentage points more other routine-labor results in a 2.19 percentage points displacement.<sup>10</sup> Similarly, Autor and Dorn (2013) find a 2.95 percentage points displacement per decade in the US (see Table 3 in their paper). In contrast, apprentice routine-labor has no displacement. These results are robust to a number of additional controls, e.g. the employment share in services, the female employment share, the share of immigrants, the share of youth (age 25 and below), and

<sup>9</sup>The results are available from the authors upon request.

<sup>10</sup>Using the separate routine measure, gives a slightly smaller, but still comparable impact of 1.7 percentage points in a decade.

**Table 3: Routine Displacement, Apprentices and Tasks**

|                | (1)                     | (2)                     | (3)                  | (4)                  | (5)                  |
|----------------|-------------------------|-------------------------|----------------------|----------------------|----------------------|
| VAR            | $\Delta^{(t+10)-t}$ RTI | $\Delta^{(t+10)-t}$ RTI | $\Delta^{89-79}$ RTI | $\Delta^{99-89}$ RTI | $\Delta^{08-99}$ RTI |
| RTI            | -0.0710<br>(0.0400)     |                         |                      |                      |                      |
| ARTI           |                         | 0.010<br>(0.130)        | -0.003<br>(0.237)    | 0.064<br>(0.100)     | 0.025<br>(0.228)     |
| ORTI           |                         | -0.219***<br>(0.052)    | -0.079<br>(0.049)    | -0.406***<br>(0.120) | -0.207**<br>(0.089)  |
| N              | 546                     | 546                     | 182                  | 182                  | 182                  |
| R <sup>2</sup> | 0.309                   | 0.371                   | 0.152                | 0.490                | 0.163                |

**Notes:** Standard errors are clustered at the federal state level and displayed in parentheses. \*\*\* significant at 1 percent, \*\* significant at 5 percent, \* significant at 10 percent. All regressions include federal state fixed effects (11 federal states). Columns (1)-(2) include time fixed effects (3 decades). Columns (2)-(5) control for the region's initial high-skilled share. All models are weighted by each periods' initial employment shares.

the share of part-time workers. With these additional controls, the coefficient on apprentice routine-labor is still insignificant. The coefficient on other routine-labor varies between  $-0.395$  and  $-0.215$ , and is significant at the 1 percent confidence level. Detailed regression results are found in Tables C.1 and C.2 in Appendix C.

Columns (3) - (5) show results separately for each decade.<sup>11</sup> Splitting the effect by decade shows that most of the displacement took place for other workers between 1989 and 1999. In fact, the 1990s were the period where ICT adoption accelerated. Nordhaus (2007) shows that computer power increased and prices decreased the most between 1985 and 1995.

Table 4 shows the effect over a longer time horizon, from 1979 to 1999 and from 1979 to 2008. This long-term perspective is done by regressing the overall change in routine employment shares between  $t = 0$  and  $t = T$  on the initial routine intensity in 1979. In the long-run, aggregate effects are negative and significant

**Table 4: Routine Displacement, Apprentices and Tasks in the Long-run**

|                | (1)                  | (2)                  | (3)                  | (4)                  |
|----------------|----------------------|----------------------|----------------------|----------------------|
| VARIABLES      | $\Delta^{99-79}$ RTI | $\Delta^{99-79}$ RTI | $\Delta^{08-79}$ RTI | $\Delta^{08-79}$ RTI |
| RTI            | -0.182*<br>(0.0921)  |                      | -0.124*<br>(0.0652)  |                      |
| ARTI           |                      | 0.036<br>(0.209)     |                      | 0.023<br>(0.204)     |
| ORTI           |                      | -0.472**<br>(0.142)  |                      | -0.448***<br>(0.046) |
| N              | 182                  | 182                  | 182                  | 182                  |
| R <sup>2</sup> | 0.318                | 0.484                | 0.342                | 0.535                |

**Notes:** Standard errors are clustered at the federal state level and displayed in parentheses. \*\*\* significant at 1 percent, \*\* significant at 5 percent, \* significant at 10 percent. All regressions include federal state fixed effects (11 federal states). Columns (2)-(4) control for the region's initial high-skilled share in 1979. All models are weighted by each periods' initial employment shares.

<sup>11</sup>To make results comparable across decades, column (5) is adjusted for the missing years. I.e. all results in Table 3 can be interpreted as 10-year changes.

at the 10 percent level. In terms of magnitude, the coefficient for the first twenty years suggests that a region with a 10 percentage points higher routine labor share has a 0.9 percentage points lower routine-labor share within 10 years, and an 1.8 percentage point decrease after 20 years. Over the entire 29 years, the result is slightly smaller but still similar. The results for columns (2) and (4) are identical to above. There is no effect for apprentice routine-labor. For other routine labor the effect is similar to the US. Comparing the results from Tables 3, 4 and Figure 7 suggest that the displacement results are robust to different specifications.

### C. Routine Shares and Service Employment

Having determined the displacement of routine-labor, we need to establish whether this displacement comes with a rise in low-skilled services. The low-skill service sector is defined as the share of regional employment that works in certain service occupations. Following the definition of Blossfeld (1985), low-skilled service occupations range from “hairstylist” and “street and indoor cleaners” to “attending on guests” and “nursing assistants.”

Table 5 regresses the change in the low-skilled service sector share on the same set of variables as above. The results are presented in decades, similar to regressions (19) and (20). Given the cross derivative, equation (15), the only difference is that we expect the coefficients to have the opposite sign.

**Table 5: Low-Skilled Services, Apprentices and Tasks**

| VARIABLES      | (1)<br>$\Delta^{(t+10)-t}$ LS | (2)<br>$\Delta^{(t+10)-t}$ LS | (3)<br>$\Delta^{(t+10)-t}$ LS | (4)<br>$\Delta^{99-79}$ LS | (5)<br>$\Delta^{08-79}$ LS |
|----------------|-------------------------------|-------------------------------|-------------------------------|----------------------------|----------------------------|
| RTI            | 0.023**<br>(0.009)            |                               |                               |                            |                            |
| ARTI           |                               | -0.014<br>(0.0213)            | -0.008<br>(0.069)             | -0.074<br>(0.093)          | -0.178<br>(0.159)          |
| ORTI           |                               | 0.024<br>(0.021)              | 0.028**<br>(0.012)            | 0.072***<br>(0.020)        | 0.111***<br>(0.032)        |
| N              | 546                           | 546                           | 546                           | 182                        | 182                        |
| R <sup>2</sup> | 0.499                         | 0.499                         | 0.499                         | 0.209                      | 0.214                      |

**Notes:** Standard errors are clustered at the federal state level and displayed in parentheses. \*\*\* significant at 1 percent, \*\* significant at 5 percent, \* significant at 10 percent. All regressions include federal state fixed effects (11 federal states). Columns (1)-(3) include time fixed effects (3 decades). Column (2) controls for the region's initial high-skilled share. All models are weighted by each periods' initial employment shares. The regional low-skill service share (dependent variable) is defined as the employment share working in low-service occupations defined after Blossfeld (1985).

Column (1) suggests that routine-labor and growth in low-skilled services are related. In magnitude, this is about one-third of the US effect. Given the lack of polarization at the lower end of the wage distribution (see Figure 1), it is to be expected that the results for the share of low-skilled services are smaller in magnitude. Column (2) controls for the high-skilled labor share and splits the results by apprentices versus other routine-labor. Both results are insignificant when controlling for the high-skilled labor share. However, consistent with the theory, the effect on other routine-labor is larger and positive (and almost significant). As it turns out, not controlling for the high-skilled share, column (3), preserves the same results, but now the coefficient on other routine-labor is significant at the 5 percent level. This result is consistent with Manning (2004) and Mazzolari and Ragusa (2013), who

**Figure 8: Wage Growth and 1979 Wage Distribution**

suggest that high-skilled labor demands low-skilled services. Thus, the more high-skilled labor the higher the demand for low-skilled services. The last two columns, (4) and (5), show the long-run trends. The results are similar in magnitude for the entire time period. Having a 10 percentage points larger other routine labor share in 1979 leads to a growth in low-skilled services of 1.1 percentage points over the entire period, compared to a 1.1 percentage point over one decade in the US (see Autor and Dorn, 2013, Table 5).

#### *D. Wage Polarisation*

Unlike the US (see Figure 1 panel B in Autor and Dorn, 2013), Germany has seen little *wage* polarisation in terms of wage growth along the skill distribution (see Figure 8).

Consistent with the complementarity of ICT technologies and high-skilled labor, the top of the skill distribution has had the highest wage growth. Dustmann et al. (2009) find a rise in the wage differential of middle-skilled (apprenticeship holders) relative to the low-skilled starting in the 1980s. However, they find no clear trend between the high- and middle-skilled.

In addition, the literature has found little wage polarization in Germany, but rather wage dispersion (consistent with employment polarization). The conclusion is that wage adjustments have been avoided due to institutional policies, e.g. centralized wage bargaining and generous unemployment benefits (see Dustmann et al., 2009; Kohn, 2006; Antonczyk et al., 2010; Senftleben and Wielandt, 2012). Given the limited evidence in our data sample, and that an analysis of wage polarization would need to account for different regional institutional factors, it is outside the scope of this study to further investigate these trends.

## V. Conclusion

ICT technology adoption has led to substantial employment polarization across the world. The standard theory suggests that routine tasks performed by the middle-skilled are most prone to displacement by computers. However, when quantifying the investment in computer adoption over the last decades across German regions, the data suggests that regions with the least routine employment have the most computer adoption. This apparent puzzle - which stands in sharp contrast to the US - is resolved by studying the composition of non-college labor in Germany.

In this paper, we develop a stylized one-sector model that demonstrates the importance of the educational-system, i.e. general versus specific training, in incentivizing firms to adopt skill-replacing technologies. Since firms that train apprentices invest resources in doing so, the adoption of ICT technologies that replace non-college labor is more costly. As apprentices are more productive than other middle-skilled labor, due to their specific training, regions with a large number of apprentices see less ICT adoption, but also less displacement of routine-labor.

For the empirical analysis, we utilize regional variation in apprentice-intensity across German local labor markets prior to the 1980s. The empirical results show virtually no displacement of apprentice routine-labor. In terms of routine-labor displacement of other (non-apprentice) workers the effects is similar to the US results. The results on computer adoption and service employment are in the same direction as in the US, but magnitudes are smaller.

The results linking employment polarization and a skill-specific labor force suggests potential policy avenues in tackling inequality. Nonetheless, the effects of slower technology adoption on growth and structural transformation should not be ignored, especially as women mostly work service occupations and apprentice employment is focused on male dominated industry occupations.

## References

- Acemoglu, Daron, and David H. Autor.** 2011. *Skills, Tasks and Technologies: Implications for Employment and Earnings*, volume 4 of *Handbook of Labor Economics*, chapter 12, (pp. 1043–1171). Elsevier.
- Acemoglu, Daron, and Jörn-Steffen Pischke.** 1998. Why Do Firms Train? Theory And Evidence. *The Quarterly Journal of Economics*, 113(1), 78–118.
- Acemoglu, Daron and Fabrizio Zilibotti.** 2001. Productivity Differences. *The Quarterly Journal of Economics*, 116(2), 563–606.
- Adda, Jerome , Christian Dustmann, Costas Meghir, and Jean-Marc Robin.** 2006. Career Progression and Formal versus On-the-Job Training. IZA Discussion Papers 2260, Institute for the Study of Labor (IZA).
- Antonczyk, Dirk, Thomas DeLeire, and Bernd Fitzenberger.** 2010. Polarization and Rising Wage Inequality: Comparing the U.S. and Germany. IZA Discussion Papers 4842, Institute for the Study of Labor (IZA).
- Autor, David. H. and David Dorn.** 2013. The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5), 1553–97.
- Autor, David H., Frank Levy, and Richard J. Murnane.** 2003. The Skill Content Of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4), 1279–1333.
- BiBB.** 1977. *Berufsbildungsbericht*. Ab 2000: BMBF Publik. Bonn: Bundesminister fuer Bildung und Wissenschaft.

- BiBB.** 2011. *Berufsbildungsbericht*. Ab 2000: BMBF Publik. Bonn: Bundesminister fuer Bildung und Wissenschaft.
- Blossfeld, Hans-Peter.** 1985. *Bildungsexpansion und Berufschancen: Empirische Analysen zur Lage der Berufsanfänger in der Bundesrepublik*. Frankfurt: Campus.
- Casey, Bernard.** 1991. Recent Developments in the German Apprenticeship System. *British Journal of Industrial Relations*, 29(2), 205–222.
- Dorner, Matthias, Marion Koenig, and Stefan Seth.** 2011. Sample of Integrated Labour Market Biographies. Regional File 1975-2008 (SIAB-R 7508). FDZ Datenreport. Documentation on Labour Market Data 201107, Institut fuer Arbeitsmarkt- und Berufsforschung (IAB), Nuernberg [Institute for Employment Research, Nuremberg, Germany].
- Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg.** (2009). Revisiting the German Wage Structure. *The Quarterly Journal of Economics*, 124(2), 843–881.
- Dustmann, Christian, and Sonia C. Pereira.** 2008. Wage Growth and Job Mobility in the United Kingdom and Germany. *Industrial and Labor Relations Review*, 61(3), 374–393.
- Eckey, Hans-Friedrich.** 1988. Abgrenzung Regionaler Arbeitsmrkte. *Raumforschung und Raumordnung*, 1(2), 24–33.
- Eckey, Hans-Friedrich, and Paul Klemmer.** (1991). Neuabgrenzung von Arbeitsmarktregionen fr die Zwecke der Regionalen Wirtschaftspolitik. *Informationen zur Raumentwicklung*, 9(10), 569–577.
- Fitzenberger, Bernd, Aderonke Osikominu, and Robert Vlter.** (2006). Imputation Rules to Improve the Education Variable in the IAB Employment Subsample. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift fr Wirtschafts- und Sozialwissenschaften*, 126(3), 405–436.
- Goldin, Claudia D., and Lawrence F. Katz.** 2008. *The Race between Education and Technology*. Belknap Press of Harvard University Press, Cambridge, Mass.
- Goos, Maarten, Alan Manning, and Anna Salomons.** 2011. Explaining job polarization: the roles of technology, offshoring and institutions. Technical report, Katholieke Universiteit Leuven.
- Hanushek, Eric A., Ludger Woessmann, and Lei Zhang.** 2011. General Education, Vocational Education, and Labor-Market Outcomes over the Life-Cycle. NBER Working Papers 17504, National Bureau of Economic Research.
- Harhoff, Dietmar, and Kane, Thomas J.** 1993. Financing Apprenticeship Training: Evidence from Germany. NBER Working Papers 4557, National Bureau of Economic Research.
- Kohn, Karsten.** 2006. Rising Wage Dispersion, After All ! The German Wage Structure at the Turn of the Century. ZEW Discussion Papers 06-31, ZEW - Zentrum fr Europaeische Wirtschaftsforschung / Center for European Economic Research.
- Koller, Martin and Barbara Schwengler.** 2000. *Struktur und Entwicklung von Arbeitsmarkt und Einkommen in den Regionen*. Number 232 in Beitrge zur Arbeitsmarkt- und Berufsforschung. Nrnberg: Inst. fr Arbeitsmarkt- und Berufsforschung der Bundesanst. fr Arbeit.
- Krueger, Dirk and Krishna B. Kumar.** 2004. Skill-Specific rather than General Education: A Reason for US–Europe Growth Differences? *Journal of Economic Growth*, 9(2), 167–207.
- Manning, Alan.** 2004. We Can Work It Out: The Impact of Technological Change on the Demand for Low-Skill Workers. *Scottish Journal of Political Economy*, 51(5), 581–608.
- Mazzolari, Francesca and Giuseppe Ragusa.** 2013. Spillovers from High-Skill Consumption to Low-Skill Labor Markets. *The Review of Economics and Statistics*, 95(1), 74–86.
- Michaels, Guy, Ashwini Natraj, and John Van Reenen.** 2010. Has ICT Polarized Skill Demand? Evidence from Eleven Countries over 25 years. NBER Working Papers 16138, National Bureau of Economic Research.
- Nordhaus, William D.** 2007. Two Centuries of Productivity Growth in Computing. *The Journal of Economic History*, 67(01), 128–159.

- Senftleben, Charlotte, and Hanna Wielandt.** 2012. The Polarization of Employment in German Local Labor Markets. Number 13 in Sonderforschungsbereich 649: Oekonomisches Risiko. Wirtschaftswissenschaftliche Fakultät.
- Soskice, David.** 1994. Reconciling markets and institutions: The German apprenticeship system. In *Training and the private sector* (pp. 25–60). University of Chicago Press.
- Spitz-Oener, Alexandra.** 2006. Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, 24(2), 235–270.
- van Ark, Bart, Robert Inklaar, and Robert H. McGuckin.** 2003. The Contribution of ICT-Producing and ICT-Using Industries to Productivity Growth: A Comparison of Canada, Europe and the United States. *International Productivity Monitor*, 6, 56–63.
- Winkelmann, Rainer.** (1997). How young workers get their training: A survey of Germany versus the United States. *Journal of Population Economics*, 10(2), 159–170.

## A. MODEL APPENDIX

This Appendix provides all formal proofs for the model.

*Derivation of the Threshold Cut-off (Equation (6))*

From Equation (6), we see that the threshold value  $J^*$  is such that,

$$\left( \frac{d^2\omega(J_t)}{dJ_t^2} + \frac{1-2J}{(J_t(1-J_t))^2} \right) < 0.$$

By Assumptions 1 and 2 the first term in brackets is always negative. The critical threshold value,  $J^*$ , must satisfy,  $J^* < \frac{1}{2}$ , since for any value  $J^* \geq \frac{1}{2}$  the expression is unambiguously negative. Hence,  $J^*$  is given as the solution to ,

$$-\frac{d^2\omega(J^*)}{d(J^*)^2} [(J^*)^4 - 2(J^*)^3 + (J^*)^2] + 2J^* - 1 = 0.$$

Above this threshold value, technological progress leads to a larger fall in the threshold, the higher the initial threshold  $J_t$  is.

*Proof of the Apprentice Productivity Schedule (Condition (14))*

PROOF:

The differential is,

$$\begin{aligned} \frac{d \frac{dJ_t}{d \ln(N_t)}}{d\lambda} &= \frac{\partial \frac{dJ_t}{d \ln(N_t)}}{\partial J_t} \frac{dJ_t}{d\lambda} + \frac{\partial \frac{dJ_t}{d \ln(N_t)}}{\partial \lambda} \\ &= \frac{d^2 J_t}{dJ_t d \ln(N_t)} \frac{dJ_t}{d\lambda} - \left( \frac{dJ_t}{d \ln(N_t)} \right)^2 \frac{d^2 \omega(J_t)}{dJ_t d\lambda} \\ &= - \left( \frac{dJ_t}{d \ln(N_t)} \right)^2 \left[ \left( \frac{d^2 \omega(J_t)}{dJ_t^2} + \frac{1-2J}{(J_t(1-J_t))^2} \right) \left( \frac{-1}{\hat{\alpha}_\ell(J_t)} \frac{d\hat{\alpha}_\ell(J_t)}{d\lambda} \frac{dJ_t}{d \ln(N_t)} \right) + \frac{d^2 \omega(J_t)}{dJ_t d\lambda} \right]. \end{aligned}$$

The final equality follows from Equations (6) and (12). Simplifying this expression results in Equation (14).

*Proof of Apprentice Productivity (Equation (15))*

There are three cross derivatives: (1) machine adoption, (2) routine displacement, and (3) low-skilled service growth.

## MACHINE ADOPTION ACROSS REGIONS

PROOF:

The proof of the machine displacement across regions can be done separately for each term. From Equation (9), the first term is,

$$\begin{aligned} \frac{\partial \left( \frac{\partial \ln(X_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} \right)}{\partial J_t} &= \left[ \frac{dJ_t}{d \ln(N_t)} \left( \frac{1}{\bar{x} - J_t} + \frac{1}{1 - J_t} \right) - \frac{d^2 J_t}{d \ln(N_t) dJ_t} \right] \\ &\quad \times \left( \frac{1}{\bar{x} - J_t} - \frac{1}{1 - J_t} \right) < 0. \end{aligned}$$

The first term in brackets is negative from equation (5) and Assumption 1. The second term is positive from equation (6). Consequently, as long as  $J_t > J^*$ , technical progress in apprentice-regions leads to less machine adoption as the threshold drops at a slower rate.<sup>12</sup> In addition, from Equation (12), we know that,

$$\frac{dJ_t}{d\lambda} > 0,$$

making the indirect effect negative for all  $J_t > J^*$ . The second term of Equation (15), the direct effect, is

$$\frac{\partial \left( \frac{\partial \ln(X_t)}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} \right)}{\partial \lambda} = - \left( \frac{1}{\bar{x} - J_t} - \frac{1}{1 - J_t} \right) \left( \frac{dJ_t}{d \ln(N_t)} \right)^2 \left( - \frac{\partial^2 \hat{\omega}(J_t)}{\partial J_t \partial \lambda} \right).$$

The sign of this expression depends on the exact  $\hat{\alpha}_\ell(i)$ -slope. Collecting terms, for Equation (15) to be negative, it must be that

$$\left[ \left( \frac{1}{\bar{x} - J_t} + \frac{1}{1 - J_t} \right) + \left( \frac{d^2 \hat{\omega}(J_t)}{dJ_t^2} + \frac{1 - 2J}{(J_t(1 - J_t))^2} \right) \frac{dJ_t}{d \ln(N_t)} \right] \frac{1}{\hat{\alpha}_\ell(i)} \frac{d\hat{\alpha}_\ell(J_t)}{d\lambda} > \frac{\partial^2 \hat{\omega}(J_t)}{\partial J_t \partial \lambda},$$

which holds under Condition (14).

#### ROUTINE DISPLACEMENT ACROSS REGIONS

PROOF:

The prove of routine displacement is analogous to machine displacement. That is, the first term from Equation (10), is

$$\frac{\partial \left( \frac{\partial \ln L_{RT,t}}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} \right)}{\partial J_t} = \frac{-\underline{x}(2J_t - \underline{x})}{J_t^2(J_t - \underline{x})^2} \frac{dJ_t}{d \ln(N_t)} + \frac{\underline{x}}{J_t(J_t - \underline{x})} \frac{d^2 J_t}{d \ln(N_t) dJ_t} > 0.$$

The second effect is negative and the first is positive for all  $J_t > J^*$ . Equation (12),

$$\frac{dJ_t}{d\lambda} > 0,$$

makes the first term unambiguously positive. The direct effect on routine displacement is,

$$\frac{\partial \left( \frac{\partial \ln L_{RT,t}}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} \right)}{\partial \lambda} = \frac{\underline{x}}{J_t(J_t - \underline{x})} \left( \frac{dJ_t}{d \ln(N_t)} \right)^2 \left( - \frac{d^2 \hat{\omega}(J_t)}{dJ_t d\lambda} \right).$$

The algebraic sign depends on the exact  $\hat{\alpha}_\ell(i)$ -slope. Collecting terms, for Equation (15) to be positive, it must be that

$$\left[ \left( \frac{(2J_t - \underline{x})}{J_t(J_t - \underline{x})} \right) + \left( \frac{d^2 \hat{\omega}(J_t)}{dJ_t^2} + \frac{1 - 2J}{(J_t(1 - J_t))^2} \right) \frac{dJ_t}{d \ln(N_t)} \right] \frac{1}{\hat{\alpha}_\ell(i)} \frac{d\hat{\alpha}_\ell(J_t)}{d\lambda} > \frac{d^2 \hat{\omega}(J_t)}{dJ_t d\lambda},$$

<sup>12</sup>Technically, this still holds for thresholds below  $J^*$ , as long as the absolute value of the first term is larger than the second term.

which is always true under Condition (14).

#### LOW-SKILLED SERVICE GROWTH ACROSS REGIONS

PROOF:

The first term of Equation (15) is,

$$\frac{\partial \left( \frac{\partial \ln L_{LST,t}}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} \right)}{\partial J_t} = \frac{1}{J_t^2} \frac{dJ_t}{d \ln(N_t)} - \frac{1}{J_t} \frac{d^2 J_t}{d \ln(N_t) d J_t} < 0.$$

The first effect is negative and the second is positive for all  $J_t > J^*$ . Since, from Equation (12),

$$\frac{dJ_t}{d\lambda} > 0,$$

the first term of Equation (15) is negative. The direct effect is,

$$\frac{\partial \left( \frac{\partial \ln L_{LST,t}}{\partial J_t} \frac{dJ_t}{d \ln(N_t)} \right)}{\partial \lambda} = -\frac{1}{J_t} \left( \frac{dJ_t}{d \ln(N_t)} \right)^2 \left( -\frac{d^2 \hat{\omega}(J_t)}{d J_t d \lambda} \right).$$

The algebraic sign depends on the exact  $\hat{\alpha}_\ell(i)$ -slope. Collecting terms, for Equation (15) to be negative, it must be that

$$\left[ \left( \frac{1}{J_t} \right) + \left( \frac{d^2 \hat{\omega}(J_t)}{d J_t^2} + \frac{1 - 2J}{(J_t(1 - J_t))^2} \right) \frac{dJ_t}{d \ln(N_t)} \right] \frac{1}{\hat{\alpha}_\ell(i)} \frac{d \hat{\alpha}_\ell(J_t)}{d \lambda} > \frac{d^2 \hat{\omega}(J_t)}{d J_t d \lambda},$$

which is always true under Condition (14).

#### B. DATA APPENDIX

Two main data sources are used: (1) the Sample of Integrated Labour Market Biographies - Regional File 1975-2008 (SIAB); and (2) German Federal Institute for Vocational Training (BIBB) Qualification and Career Survey (QCS) 1979 and 1999.

##### B1. The SIAB Regional File 1975-2008

The SIAB is a 2% representative random sample of the German workforce collected by the Institute for Employment Research (IAB). It covers all currently employed individuals that are subject to social security payments for the years 1975 to 2008. It excludes self-employed, civil servants, individuals doing their military service and students. Marginally employed are only considered after 1999.

#### SAMPLE SELECTION AND VARIABLE DESCRIPTION

##### EMPLOYMENT.

The sample is restricted to males and females in West Germany. We drop individuals whose status of employment is coded as “doing an apprenticeship/traineeship,” “doing an internship” or those that have an undefined employment status. We have information on whether a worker works part-time (no hours) and hence follow Dustmann et al. (2009) and weight part-time workers by 2/3.

## EDUCATION.

The education variable is based on extrapolated data following imputation *method 1* in Fitzenberger et al. (2006). The high-skilled are defined as workers who graduated from university or college. Apprentices are classified as individuals that obtained an apprenticeship degree within the same broad sector (of services or non-services) they currently work in.

## WAGES.

When ranking occupations, we compute mean wages for each occupation in 1979. We only consider full-time workers for the ranking. Since wages are top-coded, we follow Dustmann et al. (2009) and impute censored wages by a fixed factor of 1.2.

## DESCRIPTIVE STATISTICS

Table B.1 shows the (unweighted) mean and the standard deviation of the main variables used. As expected, routine shares decreased over time, while abstract

**Table B.1:** Descriptive Statistics 1979 and 2008 across Regions

|                   | 1979          | 2008                       |
|-------------------|---------------|----------------------------|
| RTI Share         | 0.359 (0.061) | 0.293 (0.055)              |
| Routine Share     | 0.366 (0.064) | 0.286 (0.061)              |
| Manual Share      | 0.382 (0.055) | 0.316 (0.056)              |
| Abstract Share    | 0.340 (0.044) | 0.427 (0.047)              |
| PC Share          |               | 0.475 <sup>a</sup> (0.044) |
| Service Share     | 0.420 (0.095) | 0.602 (0.096)              |
| Low Service Share | 0.183 (0.029) | 0.221 (0.029)              |
| Apprentice Share  | 0.630 (0.043) | 0.584 (0.033)              |
| Female Share      | 0.363 (0.044) | 0.443 (0.035)              |
| Immigrant Share   | 0.072 (0.041) | 0.060 (0.030)              |
| Young Share       | 0.218 (0.038) | 0.094 (0.016)              |
| Age               | 37.1 (1.614)  | 42.5 (0.771)               |
| Part-time Share   | 0.077 (0.020) | 0.318 (0.036)              |

shares increase. The PC measure comes from the QCS 1999 (see B.B2 for details). The share of service sector employment increases, as well as the share of female workers and the share of workers in part-time work (highly correlated with female labor market participation). Average age increases due to the demographic transition, together with higher university attendance rates, it significantly decreases the share of young workers (below age 26) from 1979 to 2008.

Table B.2 shows changes in broad occupation employment shares over time. The middle of the skill distribution, *Production/Craft*, has fallen. However, unlike the US, most of the fall in production has been absorbed by a rise in professional employment and not in services.

Using the same occupation classification, Table B.3 computes average task measures of the routine, manual and abstract components. In line with the theory and empirical results, occupations with the largest employment fall are the most routine-intensive occupations. Occupations with the largest increase are most abstract.

**Table B.2:** Employment Shares by Broad Occupation Class

| Occupation Class          | Employment Shares |       | Employment Changes |                  |
|---------------------------|-------------------|-------|--------------------|------------------|
|                           | 1979              | 2008  | $\Delta_{79-08}$   | $Growth_{79-08}$ |
| Managers                  | 0.023             | 0.031 | 0.008              | 33.4%            |
| (Semi-)Professionals      | 0.043             | 0.092 | 0.049              | 113.8%           |
| Technicians/Engineers     | 0.069             | 0.074 | 0.006              | 8.1%             |
| Commercial/Administration | 0.261             | 0.302 | 0.042              | 16.0%            |
| Production/Craft          | 0.411             | 0.267 | -0.144             | -35.0%           |
| Agricultural occupations  | 0.010             | 0.012 | 0.001              | 14.0%            |
| Services                  | 0.183             | 0.222 | 0.038              | 20.9%            |

**Table B.3:** Tasks by Broad Occupation Class

| Occupation Class          | $RTI$ | Routine | Manual | Abstract |
|---------------------------|-------|---------|--------|----------|
| Managers                  | 0.379 | 4.450   | 2.523  | 6.021    |
| (Semi-)Professionals      | 0.358 | 4.318   | 3.051  | 6.125    |
| Technicians/Engineers     | 0.408 | 5.105   | 3.852  | 5.451    |
| Commercial/Administration | 0.413 | 4.678   | 2.912  | 5.196    |
| Production/Craft          | 0.477 | 6.433   | 5.903  | 4.091    |
| Agricultural occupations  | 0.407 | 4.891   | 5.002  | 4.614    |
| Services                  | 0.438 | 5.197   | 4.532  | 4.380    |

## B2. The Qualification and Career Survey

The QCS is a representative survey of employees carried out by the BiBB and the IAB. It contains four cross-sections, in 1979, 1985/86, 1991/92 and 1998/99, where each covers about 30,000 individuals. We use these datasets from 1979 and 1999 to construct the computer utilization measures and occupation-specific task measures.

### COMPUTER MEASURE

People are asked whether they use personal computers (PCs) during their regular work. We construct an occupation  $k$  specific PC measure as the average share of workers within each occupation that use a PC. Since PC usage varies between different sectors, we differentiate between the broad sector of services and non-services (sector is denoted by  $s$ , occupation by  $k$ , individual by  $i$ ):

$$PC_{sk}^{99} = \left( \sum_{s=1}^S \sum_{i=1}^I L_{isk,1999} \cdot 1 [PCuse_{isk,1999} = YES] \right) \left( \sum_{s=1}^S \sum_{i=1}^I L_{isk,1999} \right)^{-1}.$$

Using the SIAB panel and regional employment shares of each occupation in 1999 as weights, a weighted mean of PC usage within a region  $j$  is,

$$PC_j^{99} = \left( \sum_{s=1}^S \sum_{k=1}^K L_{jsk,1999} \cdot PC_{sk}^{99} \right) \left( \sum_{s=1}^S \sum_{k=1}^K L_{jsk,1999} \right)^{-1}.$$

## OCCUPATION SPECIFIC TASKS

To construct occupation-specific task intensities we use the QCS 1979 survey, since these should reflect “initial” task requirements prior to computerization (for further details see Spitz-Oener, 2006). Similar to the US DOT measures, four measures are computed from the QCS,

- 1) *Routine task*: Routine intensity measured by how often single work steps repeat - classified from 1 to 5, i.e. from very repetitive to not at all.
- 2) *Manual task*: Manual task intensity measured by the intensity of dexterity (“Handgeschick und Fingerfertigkeit”) - on a scale from 1 to 5, from “(almost) always required” to “hardly any dexterity required.”
- 3) *Non-routine interactive task*: Classified by the intensity of required planning and coordination skills - on the scale 1 to 5, from very intense in coordination requirements to not at all.
- 4) *Non-routine analytic task*: Measured by the occupations’ math requirements - classified into five categories from very basic arithmetic operations to very advanced arithmetic knowledge including differential calculus, integrals and algebra.

## B3. Further Datasources

Figure 1a for the US uses the 1980 and 2000 census. The sample includes all working individuals, that are not institutionalized, in school or in active-military duty. All observations are weighted by the US census weight multiplied by annual hours worked.

## C. EMPIRICAL APPENDIX

Below are the results for all additional robustness test of routine-labor displacement.

Table C.1: Routine Displacement and Tasks including Controls

| VAR.              | Dependent Variable RTI |                      |                   |                      |                   |                     |                     |                    |                      |
|-------------------|------------------------|----------------------|-------------------|----------------------|-------------------|---------------------|---------------------|--------------------|----------------------|
|                   | (1)                    | (2)                  | (3)               | (4)                  | (5)               | (6)                 | (7)                 | (8)                | (9)                  |
|                   | $\Delta(t+10)-t$       | $\Delta(t+10)-t$     | $\Delta(t+10)-t$  | $\Delta(t+10)-t$     | $\Delta(t+10)-t$  | $\Delta(t+10)-t$    | $\Delta(t+10)-t$    | $\Delta(t+10)-t$   | $\Delta(t+10)-t$     |
| RTI               | -0.071<br>(0.040)      | -0.188**<br>(0.062)  | -0.086<br>(0.061) | -0.165**<br>(0.054)  | -0.068<br>(0.039) | -0.107*<br>(0.054)  | -0.107**<br>(0.045) | -0.075*<br>(0.039) | -0.354***<br>(0.074) |
| Highskilled Share |                        | -0.281***<br>(0.059) |                   |                      |                   |                     |                     |                    | -0.211***<br>(0.061) |
| Lowskilled Share  |                        |                      | 0.042<br>(0.077)  |                      |                   |                     |                     |                    | 0.054<br>(0.057)     |
| Service Share     |                        |                      |                   | -0.070***<br>(0.015) |                   |                     |                     |                    | -0.117***<br>(0.019) |
| Female Share      |                        |                      |                   |                      | 0.036<br>(0.029)  |                     |                     |                    | 0.093<br>(0.051)     |
| Immigrant Share   |                        |                      |                   |                      |                   | -0.170**<br>(0.055) |                     |                    | -0.091<br>(0.049)    |
| Young Share       |                        |                      |                   |                      |                   |                     | 0.189***<br>(0.046) |                    | 0.051<br>(0.049)     |
| Part-time Share   |                        |                      |                   |                      |                   |                     |                     | 0.088<br>(0.053)   | -0.0223<br>(0.063)   |
| Observations      | 546                    | 546                  | 546               | 546                  | 546               | 546                 | 546                 | 546                | 546                  |
| R <sup>2</sup>    | 0.309                  | 0.363                | 0.310             | 0.321                | 0.310             | 0.348               | 0.337               | 0.315              | 0.403                |

**Notes:** \*\*\* significant at 1 percent, \*\* significant at 5 percent, \* significant at 10 percent.

Standard errors are clustered at the federal state level and displayed in parentheses.

All regressions include federal state fixed effects (11 federal states) and time fixed effects.

All models are weighted by each periods' initial employment shares.

**Table C.2: Routine Displacement, Apprentices and Tasks including Controls**

| VAR.              | Dependent Variable RTI     |                            |                            |                            |                            |                            |                            |                            |
|-------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
|                   | (1)<br>$\Delta_{(t+10)-t}$ | (2)<br>$\Delta_{(t+10)-t}$ | (3)<br>$\Delta_{(t+10)-t}$ | (4)<br>$\Delta_{(t+10)-t}$ | (5)<br>$\Delta_{(t+10)-t}$ | (6)<br>$\Delta_{(t+10)-t}$ | (7)<br>$\Delta_{(t+10)-t}$ | (8)<br>$\Delta_{(t+10)-t}$ |
| ARTI              | 0.010<br>(0.130)           | 0.019<br>(0.134)           | -0.111<br>(0.150)          | 0.011<br>(0.129)           | 0.036<br>(0.116)           | -0.001<br>(0.124)          | 0.011<br>(0.132)           | -0.131<br>(0.104)          |
| ORTI              | -0.219***<br>(0.052)       | -0.254***<br>(0.057)       | -0.331***<br>(0.070)       | -0.221***<br>(0.054)       | -0.215***<br>(0.052)       | -0.224***<br>(0.051)       | -0.227***<br>(0.056)       | -0.395***<br>(0.065)       |
| Highskilled Share | -0.271***<br>(0.057)       | -0.273***<br>(0.052)       | -0.283***<br>(0.060)       | -0.272***<br>(0.060)       | -0.200***<br>(0.039)       | -0.232***<br>(0.046)       | -0.290***<br>(0.069)       | -0.203***<br>(0.055)       |
| Lowskilled Share  |                            | 0.080<br>(0.051)           |                            |                            |                            |                            |                            | 0.089*<br>(0.047)          |
| Service Share     |                            |                            | -0.080***<br>(0.017)       |                            |                            |                            |                            | -0.109***<br>(0.016)       |
| Female Share      |                            |                            |                            | -0.00771<br>(0.0289)       |                            |                            |                            | 0.0691<br>(0.0413)         |
| Immigrant Share   |                            |                            |                            |                            | -0.103*<br>(0.046)         |                            |                            | -0.106*<br>(0.047)         |
| Young Share       |                            |                            |                            |                            |                            | 0.117***<br>(0.032)        |                            | 0.034<br>(0.046)           |
| Part-time Share   |                            |                            |                            |                            |                            |                            | -0.043<br>(0.060)          | -0.026<br>(0.063)          |
| Observations      | 546                        | 546                        | 546                        | 546                        | 546                        | 546                        | 546                        | 546                        |
| R <sup>2</sup>    | 0.371                      | 0.375                      | 0.386                      | 0.371                      | 0.382                      | 0.380                      | 0.372                      | 0.411                      |

**Notes:** \*\*\* significant at 1 percent, \*\* significant at 5 percent, \* significant at 10 percent. Standard errors are clustered at the federal state level and displayed in parentheses. All regressions include federal state fixed effects (11 federal states) and time fixed effects. All models are weighted by each periods' initial employment shares.

UBS International Center of Economics in Society  
University of Zurich  
Department of Economics  
Schönberggasse 1  
CH-8001 Zurich  
Tel. +41 44 634 57 22  
[contact@ubscenter.uzh.ch](mailto:contact@ubscenter.uzh.ch)  
[www.ubscenter.uzh.ch](http://www.ubscenter.uzh.ch)

*Economics. For Society.*

