I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11220

The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change

Michael J. Böhm

DECEMBER 2017

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11220

The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change

Michael J. Böhm
Vancouver School of Economics, University of British Columbia, University of Bonn and IZA

DECEMBER 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change*

The debate about the impact of routine-biased technical change on wages revolves around the question whether occupational or overall wage distributions polarized. This paper instead argues that routine task prices should decline compared to abstract and manual task prices. I propose a new method, which exploits the sorting of workers into tasks and their associated wage growth in the Roy model, to estimate the changes in task prices under relatively weak assumptions. Empirical results for male workers in two U.S. datasets indicate that task prices polarized during the 1990s and 2000s. The estimates match the upper part of the wage distribution, and they are consistent with differences across countries and time periods in the lower part.

JEL Classification: J23, J24, J31

Keywords: task prices, Roy Model, routine-biased technical change, polarization, wage distribution

Corresponding author:
Michael J. Böhm
Institute for Applied Microeconomics
University of Bonn
Adenauerallee 24 - 42
53113 Bonn

E-mail: mboehm1@uni-bonn.de

* Earlier and substantially different versions of this research were circulated as "The Wage Effects of Job Polarization: Evidence from the Allocation of Talents" and as "Has Job Polarization Squeezed the Middle Class? Evidence from the Allocation of Talents". I would like to thank Daron Acemoglu, David Autor, David Dorn, Philipp Eisenhauer, Luis Garicano, Hans-Martin von Gaudecker, Georg Graetz, David Green, Thomas Lemieux, Alan Manning, Steve Pischke, Yona Rubinstein, and Christian Siegel as well as participants at various seminars for very helpful comments.

1 Introduction

A large literature has examined the effect of routine-biased technical change ("RBTC") on the labor market. In particular, several papers have shown that during the 1990s and 2000s employment in most developed countries polarized away from middle-skill occupations that are intensive in routine tasks and into high-skill and low-skill occupations that are intensive in abstract and manual tasks ("job polarization", e.g., Autor, Levy, and Murnane, 2003; Goos and Manning, 2007; Acemoglu and Autor, 2011; Goos, Manning, and Salomons, 2014). Moreover, in the United States, wages at the top and the bottom quantiles of the (occupational) wage distribution increased while wages at the middle of the (occupational) wage distribution stagnated ("wage polarization", e.g., Autor, Katz, and Kearney, 2008; Acemoglu and Autor, 2011; Autor and Dorn, 2013).

However, recent evidence suggests that the polarization of wages may not be pervasive, casting some doubts on RBTC and its importance for wage inequality. First, wage polarization in the United States only clearly occurred during the 1990s. In the 1980s wage inequality rose across the board and in the 2000s only wages at the top of the distribution increased while the lower half largely stagnated (Acemoglu and Autor, 2011; Autor, 2015). Moreover, average wages in some of the contracting routine occupations increased and they declined in some of the expanding manual occupations (e.g., Mishel, Shierholz, and Schmitt, 2013). Finally, despite pervasive and substantial job polarization in Europe and Canada, there exists little evidence that the wage distribution has been polarizing in these countries (Dustmann, Ludsteck, and Schönberg, 2009; Naticchioni, Ragusa, and Massari, 2014; Green and Sand, 2015).

This paper shows how these seemingly contradicting facts may be reconciled. Like the preceding literature, I use a Roy (1951) model to specify labor supply in the context of RBTC. I demonstrate using simulations that different restrictions on the dependence structure of workers' skills lead to occupational and overall wage distributions which are consistent with any of the previous findings. Instead, RBTC's main prediction is about the prices that are paid for skills in tasks:[1] the routine task price will decline compared to the abstract and the manual task price ("task price polarization"). From this, an additional robust prediction emerges on the worker level: under any arbitrary assumption about the dependence structure of skills, workers' wage growth will be higher when they start out

[1] In the RBTC-Roy model a worker's wage in a given task is the product of his task-specific productivity ("skill" in that task) times the prevailing equilibrium market price per unit of that task input (the "task price").

in the rising non-routine tasks than when they start in the declining routine task.

I use this prediction at the worker level to derive a new method for estimating changes in task prices in the Roy model. As in previous literature (e.g., Heckman and Sedlacek, 1985; Autor, Katz, and Kearney, 2006; Acemoglu and Autor, 2011; Firpo, Fortin, and Lemieux, 2013), the approach assumes that RBTC affects task prices but not skills in tasks. I show that, when task prices change in the Roy model, workers' log wage growth only depends on their task choices and these changing task prices, not their skills directly. I then exploit a repeated cross-section setting where the population distribution of skills, conditional on some fundamental and time-invariant talent vector x_i, is arguably constant over time ("comparability" assumption). This allows me to measure workers' expected wage growth in relation to their task choice *propensities* conditional on x_i, and therefore to estimate the changing task prices.

This new "Propensity Method" of estimating task prices carries the advantages that it makes minimal assumptions about the cross-sectional distribution of workers' skills and that it can be implemented in a simple linear wage regression ("Propensity Regression") for three or more tasks. Monte Carlo simulations show that the method recovers the actual changes in task prices successfully under different assumptions about the dependence structure of workers' skills and that it can account for some additional confounding factors that may affect workers' wages, such as changing returns to college and the minimum wage.[2] Section 3.4 provides a detailed comparison to recent alternative approaches for estimating task prices (Yamaguchi, 2012; Firpo, Fortin, and Lemieux, 2013; Gottschalk, Green, and Sand, 2015; Cortes, 2016; Yamaguchi, 2016).

To implement the Propensity Method, I construct two cross-sections of 27 year old male workers between 1984–1992 and 2007–2009 from the cohorts of the National Longitudinal Survey of Youth (NLSY79 and NLSY97). The NLSY uniquely contains early-determined, multidimensional, and time-invariant measures of worker talents, such as mathematical, verbal, and mechanical test scores and risky behaviors. These talents predict workers' task choices ("first-stage") and they arguably fulfil the comparability assumption. Further, since the Propensity Method is derived from a discrete Roy model, I merge detailed occupations into three broad groups that are intensive in their non-routine abstract, their routine, and their non-routine manual components according to the Dictionary of Occupational Titles (DOT) and the Occupational Information Network (O*NET).

[2]Identification requires an approximation of the adjustment path of workers' sorting between the initial sorting under the old task prices and the final sorting under the new task prices. As the adjustment path is bounded between the initial and final sorting, this approximation is of second order in practice.

These broad occupation groups constitute the "tasks" for which prices are estimated.

The Propensity Method for estimating task prices provides three empirical advances in the debate about the relationship between RBTC and wages. First, the main implication of RBTC, that task prices are polarizing, is now testable. The results from the NLSY indicate that RBTC had a strong impact on task prices during the joint period of the 1990s and 2000s. In particular, the relative price that is paid per unit of skill in the abstract (manual) task rose by 25 (33) log points between 1984–1992 and 2007–2009, while the absolute price paid for routine tasks declined. This suggests task price polarization during that period. Moreover, the prediction at the worker level that wages for starters in non-routine tasks should grow compared to starters in routine tasks is also testable. I find that in the NLSY, workers who according to their observable talents are more likely to work in rising non-routine tasks have higher wage growth than workers who are likely to work in declining routine tasks. This is consistent with existing worker level evidence from Cortes (2016) and Acemoglu and Autor (2011), but at the same time strengthening these results because in the NLSY workers with the same talent vector are arguably comparable over time.

The third distinct advance conferred by the Propensity Method is that one can use the estimated task prices to return to the aggregate level from which the debate about RBTC and wages started. In the NLSY, assigning every worker the price change of their tasks done in the initial period yields an increase of inequality in the upper half of the wage distribution that resembles its actual change during the 1990s and 2000s. Predicted inequality in the lower half of the wage distribution is flat, however, while actual inequality narrowed substantially (wage polarization). Adjusting for the increase in the real value of the minimum wage provides a better (but not perfect) fit for this part, too. This finding is consistent with the results from the simulations and with evidence from other countries. Both show that employment is polarizing and that inequality in the upper half of the wage distribution is increasing, but also that changes of inequality in the lower half of the distribution is indeterminate. Despite the substantial impact of RBTC, policy variables like the changing minimum wage may therefore have played a role for the differential changes across countries in the lower half of the wage distribution.

Overall, the evidence in this paper is supportive of a substantial impact of RBTC on workers' wages. As a robustness check, I employ decennial census and American Community Survey (Census/ACS) data from Acemoglu and Autor (2011), which is an attractive complement to the NLSY because it is large and thus provides greater statistical power. It also allows estimation in additional time periods and sub-periods. Using education-age-region cells in the Census/ACS as elements of the x_i vector and applying the Propensity Method yields qualitatively similar estimates to the NLSY for the changing task prices and a relatively good fit between the predicted and the actual change in the wage distribution.

The main estimates in both datasets are unaffected when changing returns to college are included, which supports RBTC over skill-biased technological change (SBTC) as alternative hypothesis, and when the rising real value of the minimum wage is accounted for. Shifting supplies of skills are also excluded as an alternative explanation because they imply the inverse relationship between changing employment and task prices than in the data. Section 4.1 discusses the implications of RBTC versus competing or complementary hypotheses including international trade and offshoring, and SBTC combined with rising demand for services.

The new Propensity Method is the other main contribution of this paper, which may be applicable beyond RBTC. In fact, RBTC is one specific hypothesis in the more general "task model" of the labor market. This model is conceptually attractive because employers reward workers' actual skills in carrying out tasks instead of observable characteristics such as education. However, skills in tasks are typically unobserved and the selection of skills will change with the task prices. Therefore, it has been difficult and often absent from prior studies to estimate the task prices.

Section 3.4 compares the Propensity Method in detail to the extant alternative approaches for estimating task prices, arguing that its advantage is to exploit a tight economic mechanism underlying the Roy model without invoking strong assumptions about unobservables. That is, those workers who choose non-routine tasks experience (relative) wage growth when (relative) non-routine task prices increase. Therefore, employing information about wage growth and choices, one can back out the changes in task prices. The (endogenous) re-sorting of workers across tasks is accounted for by using task choices before and after the task price change. The method is also easy to implement for multiple tasks and transparent about which empirical moments are used for identification. Finally, the paper shows theoretically and using simulations which predictions about wages can and cannot be derived from an unrestricted Roy model when task prices change.

The next section shows that RBTC is in fact ambiguous about occupational or overall wage distributions, but that it has clear implications for task prices and worker-level wages. Section 3 uses this finding to derive and validate the new Propensity Method for estimating task prices. Then the occupational grouping into tasks and the main NLSY estimation sample are introduced. Section 5 presents the results in this sample. Finally, robustness of the results in the alternative Census/ACS dataset is summarized and the last section concludes. Further in-depth discussion of related literature can be found in Sections 2.2 (previous studies considering RBTC and wages), 3.4 (alternative methods for estimating task prices), and 4.1 (occupational grouping and competing hypotheses to RBTC).

2 The Roy Model under RBTC

This section studies what RBTC implies on the aggregate level using a Roy model of labor supply. I show that supposed predictions of RBTC are not robust to different assumptions of how workers' skills are distributed across tasks. Instead, RBTC has unambiguous implications for individual workers' wage growth over time. A new equation linking wage growth to task choices is derived from the unrestricted Roy model.

2.1 Setup

The models in the RBTC literature have in common that labor supply is specified in a Roy framework. Therefore, suppose there are three tasks, abstract A, routine R, and manual M. Each worker i possesses log skills $s_i = \{s_{Ai}, s_{Ri}, s_{Mi}\}$ and faces log task prices $\pi_t = \{\pi_{At}, \pi_{Rt}, \pi_{Mt}\}$. Since the whole Section 2 is a comparative statics analysis, the subscript $t \epsilon [0, 1]$ should be thought of as simply indexing different task price vectors π_t for now, not time more broadly. Potential wages are

$$w_{Kit} = \pi_{Kt} + s_{Ki} \text{ with } K \epsilon \{A, R, M\} \tag{1}$$

and observed wages are

$$w_{it} = \begin{cases} \pi_{At} + s_{Ai} & \text{if } I_A(s_i, \pi_t) = 1 \\ \pi_{Rt} + s_{Ri} & \text{if } I_R(s_i, \pi_t) = 1 \\ \pi_{Mt} + s_{Mi} & \text{if } I_M(s_i, \pi_t) = 1 \end{cases} \tag{2}$$

with choice indicators

$$
\begin{aligned}
I_A(s_i, \pi_t) &= 1[\pi_{At} + s_{Ai} > \pi_{Rt} + s_{Ri},\ \pi_{At} + s_{Ai} > \pi_{Mt} + s_{Mi}] \\
I_R(s_i, \pi_t) &= 1[\pi_{At} + s_{Ai} \leq \pi_{Rt} + s_{Ri},\ \pi_{Rt} + s_{Ri} \geq \pi_{Mt} + s_{Mi}] \\
I_M(s_i, \pi_t) &= 1[\pi_{Mt} + s_{Mi} > \pi_{Rt} + s_{Ri},\ \pi_{At} + s_{Ai} \leq \pi_{Mt} + s_{Mi}].
\end{aligned} \tag{3}
$$

The RBTC literature has proposed several general equilibrium models in which computer capital is a (relative) substitute for routine tasks in the production function (e.g., Autor, Katz, and Kearney, 2006; Autor and Dorn, 2013), in which it is raising effective routine task inputs (Cortes, 2016), or in which it replaces an increasing continuum of routine tasks (Acemoglu and Autor, 2011). Together with the labor supply setup (1)–(3), and some additional restrictions in each of the papers, all of these models have in common that a rising availability of computer capital (RBTC) leads to an increase in abstract and manual non-routine compared to routine task prices ("task price polarization"):

$$
\triangle(\pi_A - \pi_R) > 0 \text{ and } \triangle(\pi_M - \pi_R) > 0, \tag{4}
$$

with $\triangle \pi_K \equiv \pi_{K1} - \pi_{K0}$. I take this as given in the current section, showing which empirical predictions on the aggregate and the worker level do and do not result from it. In fact, RBTC and task price polarization are equivalent in the model of Equations (1)–(4), as the π_{Kt} are the only channel via which RBTC influences the labor supply side. In the empirical analysis, Equation (4) will be taken as the core prediction from the RBTC model to be tested in the data.[3]

Equation (4) implies that RBTC affects wages only via market prices for tasks, not via changing skills. This assumption has been made throughout the RBTC literature (e.g., Autor, Katz, and Kearney, 2006; Acemoglu and Autor, 2011; Firpo, Fortin, and Lemieux, 2013; Autor and Dorn, 2013; Cortes, 2016) as well as in earlier work by Heckman and Sedlacek (1985). Although restrictive, it is conceptually attractive because it imposes the fundamental idea of the task model that worker characteristics are not priced directly but via the tasks that they help providing to the market. It is also critical because it generates the empirical restrictions that are necessary for estimating the task prices below.

Finally, while it is assumed not to change because of RBTC, the joint population distri-

[3]Prior literature has not emphasized (4) as the core prediction from RBTC, because it lacked a convincing method of estimating task prices. Earlier versions of this paper explicitly derived (4) from the Autor, Katz, and Kearney (2006) model with a general distribution of workers' skills. Notice also that "general equilibrium" labor supply responses, e.g., to the low-skill manual task, may attenuate the relative price $\triangle(\pi_M - \pi_R)$ but not overturn it. Otherwise the supply response would not have occurred in the first place.

bution F_{s_A,s_R,s_M} of workers' skills is otherwise left completely unrestricted. The exception are the simulations, which adopt different specifications of F_{s_A,s_R,s_M} for illustration.

2.2 RBTC Makes No Unambiguous Predictions on the Aggregate Level

It is well established that in the two-sector Roy model rising task prices in one sector need not lead to increasing wages in that sector and that self-selection may have ambiguous effects on overall wage inequality (e.g., Heckman and Honoré, 1990). By extension, these (negative) results should carry over to the three sector case. The contribution of this section is to illustrate that, even under normality, RBTC (or task prices polarization) does not imply that average wages in tasks or the overall wage distribution will polarize. Therefore, several empirical findings in the literature that seemingly contradict RBTC are in fact potentially consistent with it.

The results in the following are illustrated with simulated data using a multivariate normal distribution of log skills in tasks. Normality implies that the variances and correlations of skill are the main parameters determining sectoral and aggregate outcomes. Under a different distribution, other parameters may matter (Heckman and Honoré, 1990). Hence, this should be understood as just one specific illustration of the more general results. Table 1 reports the parameters for the simulations. The differences between the respective left and right panels are either with respect to the variances or the correlations of skills and highlighted in bold. Sketches of the analytical proofs for the results are relegated to Appendix A.

Prediction 1 (Negative). *Task price polarization has no clear implication on wages in tasks. In particular, it need not lead to the polarization of average wages in tasks.*

The intuition behind Prediction 1 is that a changing selection bias in tasks may invert the direct effect of the task prices themselves. The top row of Figure 1 illustrates this result for a case when the price of the abstract task rises more than of the manual task and the price of the routine task falls ($\triangle(\pi_A - \pi_R) = .35$, $\triangle(\pi_M - \pi_R) = .10$, $\triangle\pi_R = -.05$).[4] In Panel (a), the correlation between abstract and manual skills in the population is low ($corr(s_{Ai}, s_{Mi}) = 0$) and average wages in tasks polarize. Conversely, in Panel (b) of Figure 1, the correlation between abstract and manual skills is high ($corr(s_{Ai}, s_{Mi}) = .7$) and,

[4]The weak increase of the manual task price admittedly makes it easier for a relatively moderate selection effect to overturn it. In the theoretical models of Autor, Katz, and Kearney (2006) and Autor and Dorn (2013), wages in the routine task may also either rise or fall, because the least able routine workers leave for the manual task.

Table 1: Parameter Values for the Simulations in Figure 1

	Figure 1 (a,b) Left Panel	Figure 1 (a,b) Right Panel	Figure 1 (c,d) Left Panel	Figure 1 (c,d) Right Panel
$\triangle(\pi_A - \pi_R)$	0.35	0.35	0.35	0.35
$\triangle(\pi_M - \pi_R)$	0.10	0.10	0.30	0.30
$\triangle\pi_R$	-0.05	-0.05	-0.20	-0.20
$var(s_{Ai})$	3.0	3.0	3.0	3.0
$var(s_{Ri})$	1.5	1.5	**1.7**	**1.1**
$var(s_{Mi})$	1.3	1.3	1.0	1.0
$corr(s_{Ai}, s_{Ri})$	0.3	0.3	0.5	0.5
$corr(s_{Ai}, s_{Mi})$	**0.0**	**0.7**	0.5	0.5
$corr(s_{Ri}, s_{Mi})$	0.0	0.0	0.5	0.5

Notes: Skills are multivariate normal with mean zero. Variances and covariances are given in the table together with the task price changes. The parameter values that differ between the respective left and right panels are emphasized in bold. $N = 10,000$ observations were drawn for each panel.

rather than polarizing, average wages in manual tasks fall even more than average wages in routine tasks. The reason is that with a high correlation of abstract and manual skills, low-skill routine workers move into the manual task while high-skill manual workers move out into abstract tasks. The selection effect then dominates the price effect.

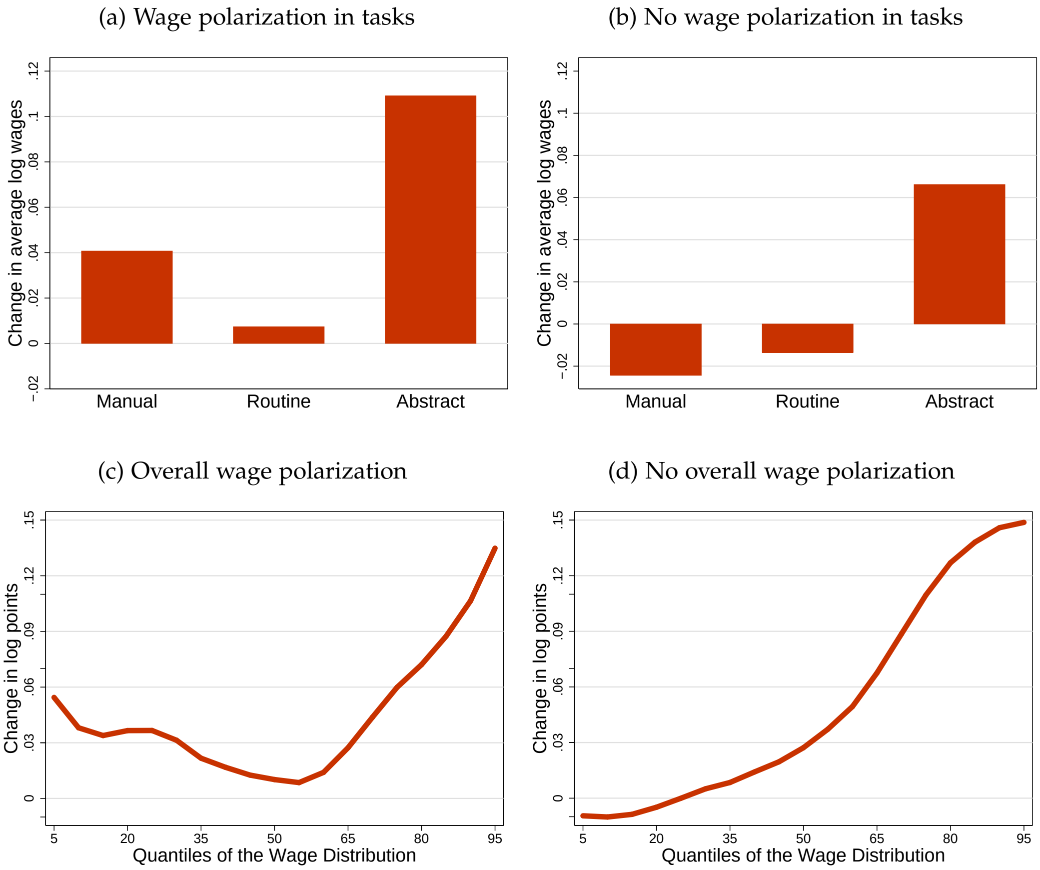
The result that RBTC need not lead to polarization of average wages in tasks is consistent with several empirical findings in the literature. In particular, during 1999–2007 employment in low-skill (service/manual-task-intensive) occupations rose strongly and at the same time wages dropped (Autor, 2015, Figures 2–4 and 6–7). Moreover, Autor and Dorn (2013) find that whereas employment contracted in routine-task-intensive clerical and sales occupations over 1980–2005, wages in these occupations increased. Mishel, Shierholz, and Schmitt (2013, p.5) also conclude from their analysis that there is "little or no connection between decadal changes in occupational employment shares and occupational wage growth" in the U.S. over the last decades. Finally, in an international context, employment in low-skill occupations increased in the UK and Canada, while at the same time wages in these occupations dropped compared to routine occupations (Goos and Manning, 2007; Green and Sand, 2015).[5]

Prediction 2 (Negative). *Task price polarization does not imply (overall) wage polarization.*

The idea behind Prediction 2 is that even if manual task workers are on average

[5] Adermon and Gustavsson (2015) find that RBTC cannot explain changing wages between occupations, but that it does have explanatory power for wages within occupations in Sweden.

Figure 1: Small differences in skill distributions can lead to qualitatively different outcomes, even under the same task price changes and a multivariate normal distribution of skills.



located at lower quantiles of the wage distribution than routine workers, it does not mean that these lower quantiles will rise more than the routine workers' quantiles. First, this is because, in the actual data, routine workers are also strongly represented in the lower end of the wage distribution (see figure 6 below). Second, with RBTC some of the manual task workers will move up in the wage distribution and overtake some of the routine task workers. Thus, not only manual task workers' initial quantiles will rise, but also the quantiles where they end up in (and vice versa for the routine task workers).[6] Empirically, this "overtaking effect" is to a greater or lesser degree always part of a change in the overall wage distribution. However, it is often assumed away in theoretical models by making workers' skill ranking one-dimensional. Such a restriction implies that wage

[6]Rising abstract task prices have a compounding effect for inequality at the top of the wage distribution. They raise abstract workers' already high initial quantiles as well as the even higher quantiles that these workers end up in.

polarization immediately follows from task price polarization.[7]

Generally this is not the case, which is illustrated in the second row of Figure 1. The relative increase in the manual task price is now assumed to be substantially higher ($\triangle(\pi_M - \pi_R) = .30$) than above ($\triangle(\pi_M - \pi_R) = .10$), since otherwise it would be hard to generate any wage polarization at all due to overtaking. In Panel (c) the variance of the routine skill in the population is high ($var(s_{Ri}) = 1.7$), which leads to a relatively large difference in initial wages between routine and manual workers and thus little overtaking when task prices change. Wages at the lowest quantiles of the wage distribution therefore increase compared to the quantiles located toward the middle. In Panel (d), because of lower routine-skill variance ($var(s_{Ri}) = 1.1$), initial wage differences between routine and manual workers are not as large. This leads to substantial overtaking when task prices change and an increase in wage inequality across-the-board instead of wage polarization. Hence, even when the task price changes are the same, one may obtain overall wage polarization or not with just a small modification of the skill distribution.

The result that RBTC may or may not lead to wage polarization is consistent with several empirical findings in the literature. Both employment and the wage distribution polarized in the United States over the 1990s and early 2000s. However, only employment in manual tasks expanded in the subperiod of the early 2000s, while relative wages only increased at the top of the wage distribution compared to the middle during that period (e.g., Acemoglu and Autor, 2011, Figures 7–10; Autor 2015). In addition, as mentioned above, a couple of recent papers find that job polarization already started in the 1980s in the United States (Mishel, Shierholz, and Schmitt, 2013; Bárány and Siegel, 2017), although we know that wage inequality rose firmly across-the-board during this period (as in Panel (d) of Figure 1). Finally, and probably most importantly, there is strong evidence for job polarization in other advanced countries, while there exists hardly any evidence for overall wage polarization in those countries.[8]

[7] For example, Acemoglu and Autor (2011) assume a fixed ranking of skill between individuals whereby high-skill workers have an absolute advantage in all task over middle-skill workers who in turn have an absolute advantage in all task over low-skill workers. Cortes (2016) makes a related assumption with a continuous distribution of skill. Focusing on the lower half of the wage distribution, Autor, Katz, and Kearney (2006) and Autor and Dorn (2013) assume that high-school (or low-skill) workers all have homogenous skills in the manual task and thus are ranked one-dimensionally by their heterogenous skills in the routine task. In none of these papers, by assumption, can a worker who initially earned less than another worker overtake that latter worker in the wage distribution when the relative price of the task that he has a comparative advantage in rises.

[8] For example, Dustmann, Ludsteck, and Schönberg (2009), Green and Sand (2015), Goos, Manning, and Salomons (2009), and Naticchioni, Ragusa, and Massari (2014) document job polarization for Germany, Canada, and across European countries, while Dustmann, Ludsteck, and Schönberg (2009), Card, Heining, and Kline (2013), Green and Sand (2015), and Naticchioni, Ragusa, and Massari (2014) find an increase in

Earlier versions of this paper showed that even job polarization does not need to occur under RBTC. In fact, employment in only either abstract or manual tasks has to rise relative to routine tasks. The assumptions on the skill distribution to achieve this (some strong negative correlations between skills) are unrealistic, however, and job polarization is indeed a robust empirical fact across countries and time periods. Therefore, this more theoretical possibility is omitted from the current draft. Conversely, as seen in Figure 1, it is relatively straightforward to generate a host of changes in average wages across tasks and the overall wage distribution. Many different parameter combinations are possible with three tasks, normality, and fixed task prices already, while it is ex ante not at all clear how skills in the population are distributed and that equilibrium task prices will be the same under different distributions.

Finally, it is also important to note that the argument made in Prediction 1 and 2 and in the respective simulation illustrations does not imply that differences in the effect of RBTC on the labor market need to be explained by differences in workers' skill endowments across countries and points in time. Task prices are an equilibrium outcome that depends on the interaction between production technologies, the extent and advancement of RBTC, and the skill distributions. All of these may differ across locations and will differ across time, and as one can verify in the simulated data, even small variations in these variables may lead to large differences in employment, wages, and the task prices themselves. What is to be learned from Predictions 1 and 2 is therefore that RBTC and task price polarization are in principle consistent with a host of aggregate outcomes in labor markets over the last decades, while task price polarization itself is an implication that appears in all models of RBTC that have been proposed to date. The next section derives further robust implications at the individual level which follow from this.

2.3 RBTC Does Make Unambiguous Predictions at the Individual Level

This section shows that the model of Equations (1)-(4) unambiguously predicts higher wage growth of workers who choose abstract and manual tasks compared to workers who choose routine tasks. This result will be the foundation of a new method for estimating task prices in the next section.

Consider Equation (2) and a marginal (log) wage change of a worker i who starts out

wage inequality across-the-board for those same countries and time periods.

in the A, R, or M task at $t = 0$ when task prices change:

$$dw_{i0} = \begin{cases} d\pi_{A0} & \text{if } I_A(s_i, \pi_0) = 1 \\ d\pi_{R0} & \text{if } I_R(s_i, \pi_0) = 1 \\ d\pi_{M0} & \text{if } I_M(s_i, \pi_0) = 1. \end{cases} \tag{5}$$

Here d denotes a marginal change of the respective variable over time. Due to the optimality of workers' initial $I_K(s_i, \pi_0)$, a marginal change in prices will only have a direct effect because workers' marginal wage changes are not affected by a potential reaction in their task choices. This is the envelope theorem, which implies that the effect on wages of a marginal change in π_{Kt}s is only the direct price effect (those workers who switch tasks on the margin do not experience a wage gain from it). In fact, at every π_t

$$dw_{it} = d\pi_{Rt} + I_A(s_i, \pi_t)d(\pi_{At} - \pi_{Rt}) + I_M(s_i, \pi_t)d(\pi_{Mt} - \pi_{Rt}). \tag{6}$$

Integrating this equation from task prices π_0 to task prices π_1 obtains the following relationship between workers' wage growth and their (changing) task choices:

$$\triangle w_i = \triangle \pi_R + \int_{\pi_{A0}-\pi_{R0}}^{\pi_{A1}-\pi_{R1}} I_A(s_i, \pi_t)d(\pi_{At} - \pi_{Rt}) + \int_{\pi_{M0}-\pi_{R0}}^{\pi_{M1}-\pi_{R1}} I_M(s_i, \pi_t)d(\pi_{Mt} - \pi_{Rt}) \tag{7}$$

Notice that the envelope theorem and the marginal result (6) are a tool to derive Equation (7), not an assumption. An alternative and instructive way of deriving (7) without invoking the envelope theorem for the simpler case of two tasks can be found in Appendix B.1. The detailed steps of the integration from (6) to (7) are in Appendix B.2. Both appendices show that Equation (7) holds for discrete changes in task prices of any size.

In Equation (7), the wage growth for worker i solely depends on his initial task choice and the change in his task choice on the adjustment path from the initial to the final prices. That is, relative skills in tasks determine the difference in workers' log wages at π_0 and π_1. Skill levels by themselves do not matter as $I_K(s_i, \pi_t)$s are really only a function of relative skills (e.g. $s_{Mi} - s_{Ri}$ and $s_{Ai} - s_{Ri}$, see Equation 3). Equation (7) therefore captures a key intuition from the Roy model: when task prices polarize, individuals who work in abstract and manual tasks gain relative to individuals who work in routine tasks. This is because the former can reap the benefits from higher comparative advantage in the rising tasks according to their skills. It is also intuitive that workers who start in abstract and manual from the beginning have the strongest comparative advantage in those tasks, but that switchers into abstract and manual have stronger comparative advantage than stayers in the routine task.

Prediction 3. *Task price polarization decreases the wages of workers who start out in the routine task compared to abstract or manual starters or both.*

For simplicity, assume that the changes in relative task prices are the same, that is, $\triangle(\pi_A - \pi_R) = \triangle(\pi_M - \pi_R) > 0$. In this case, according to Equation (7), stayers in abstract or manual tasks experience wage growth of $\triangle\pi_A = \triangle\pi_M$, as $I_A(s_i, \pi_t) = 1$ or $I_M(s_i, \pi_t) = 1$, respectively, for all $t\epsilon[0, 1]$. Wages of stayers in routine grow by only $\triangle\pi_R$, while wages of switchers from routine to abstract or manual grow within the interval $(\triangle\pi_R, \triangle\pi_A = \triangle\pi_M)$, depending on how early they find it optimal to switch. In either case, however, starters in routine tasks experience lower wage growth than starters in abstract or manual tasks. Appendix A proves Prediction 3 and shows that when relative task prices are not the same, say $\triangle\pi_A > \triangle\pi_M$, there may be (some) starters in routine tasks with higher wage growth than (some) starters in manual tasks if the former find it optimal to switch over quickly into abstract tasks (and vice versa when $\triangle\pi_A < \triangle\pi_M$).

Prediction 3 will be tested in the NLSY data below. The test is analogous to Acemoglu and Autor (2011)'s regression of different demographic groups' wage growth onto their initial probabilities to work in the three tasks, but the NLSY is a conceptually attractive complement to the data used in that paper. Prediction 3 therefore provides a theoretical foundation to Acemoglu and Autor (2011)'s intuitive regression specification and demonstrates that it does not depend on a specific assumption about the underlying population distribution of workers' skills (with the caveat that strictly-speaking only the relative wages of those with higher initial probabilities to work in either abstract or manual have to rise).

Prediction 3 also shows that the empirical tests in Cortes (2016) persist under a general distribution of workers' skills. As mentioned above, a corollary of Prediction 3 is that switchers into abstract or manual tasks experience higher wage growth than stayers in the routine tasks. Using panel data at the individual level, Cortes finds that the wage growth of stayers in the abstract and manual tasks is higher than the wage growth of stayers in routine tasks, but he also finds some evidence that wage growth of switchers out of routine tasks is higher than of stayers in routine tasks. Prediction 3 implies that these are robust empirical results in favor of RBTC at the individual level under any arbitrary distribution of workers' skills in tasks.[9]

[9] The prediction in Cortes (2016)'s model that the most and least skilled workers are leaving routine for

Finally, as far as I am aware, the relationship under changing task prices between workers' wage growth and their task affiliations for an unrestricted skill distribution (Equation 7) has not previously been derived in the Roy model.

3 The Propensity Method for Estimating Task Prices

This section derives the new Propensity Method for estimating task prices and discusses the assumptions about the data that need to hold to implement it. It conducts Monte Carlo simulations to verify that the method works and shows how some confounding factors other than RBTC that may affect wages can be accounted for. Finally, the method is compared to alternative approaches for estimating task prices used in the literature.

3.1 Derivation of the Propensity Regression

Result (7) is most helpful because it provides an approach of estimating price changes from data on workers' task choices and wages. One could go about this in panel data, exploiting the differential wage growth by task for a constant set of workers over time. However, among other challenges,[10] such an approach would critically rely on disentangling task prices from differential (and partly idiosyncratic) skill accumulation across tasks that occurs even in the absence of any other changes. This is difficult to do convincingly (also see the discussion in Section 3.4). The approach taken in this paper is instead to note that in some datasets one may observe characteristics ("talents") which make workers more or less likely to choose different tasks. In the sense of the model, workers' skills depend on these talents in different ways.

An intuitive example to think about this is Heckman and Sedlacek (1985)'s linear factor formulation of log wages:

$$w_{Kit} = \pi_{Kt} + s_{Ki} = \pi_{Kt} + \beta_K' x_i + u_{Ki}, \tag{8}$$

where $x_i = [x_{1i}, ..., x_{ji}, ..., x_{Ji}]'$ are the observed talents, β_K the corresponding linear projection coefficients, and $u_{Ki} = \delta_K' z_i$ a regression error which again depends on unobservable

the abstract and manual tasks, respectively, does not survive the general skill distribution. This is because the Roy model implies that marginal workers (switchers) are those with the lowest comparative advantage in their tasks (see also Equation 3), but in general it does not imply that comparative and absolute advantage align (in Cortes model they align perfectly).

[10] For example, the Roy model assumes perfect mobility between tasks. This is more likely fulfilled in repeated cross-sections for different cohorts of workers than in individual longitudinal data.

talents z_i and linear projection coefficients δ_K. This specific example is similar to Firpo, Fortin, and Lemieux (2013), who postulate that skills in tasks are a linear combination of characteristics, some observed, others not. Again, as in Heckman and Sedlacek (1985), Firpo, Fortin, and Lemieux (2013), and throughout the RBTC literature, the β_K and δ_K vectors are assumed task-specific but time-invariant, while the task prices are changing with RBTC over time. Although all the results in this paper hold for a general time-invariant dependency of s_{Ki} on observables x_i and unobservables z_i, specification (8) is adopted for better illustration from now on.

I now rewrite Result (7) in a way that is amenable to empirical analysis:

$$\begin{aligned} w_{i_1 1} - w_{i_0 0} = \triangle\pi_R + &\int_{\pi_{A0}-\pi_{R0}}^{\pi_{A1}-\pi_{R1}} I_A(s_{i_t}, \pi_t) d(\pi_{At} - \pi_{Rt}) + \\ &+ \int_{\pi_{M0}-\pi_{R0}}^{\pi_{M1}-\pi_{R1}} I_M(s_{i_t}, \pi_t) d(\pi_{Mt} - \pi_{Rt}), \end{aligned} \tag{7'}$$

where the time sub-index i_t indicates at which point in time individual i is observed, since Result (7) not only holds for one specific worker i over time but for all workers who possess the same skill vector s_i. Therefore, in Equation (7′) two indexes i_t and i_τ do not necessarily indicate the same worker but more generally two workers (e.g., i_t from an older cohort if $t > \tau$) with exactly the same skills $s_{i_t} = s_{i_\tau} = s_i$. Take expectations on both sides of (7) conditional on x_{i_t} to get

$$\begin{aligned} E(w_{i_1 1}|x_{i_1}) - E(w_{i_0 0}|x_{i_0}) = \triangle\pi_R + &\int_{\pi_{A0}-\pi_{R0}}^{\pi_{A1}-\pi_{R1}} p_A(x_{i_t}, \pi_t) d(\pi_{At} - \pi_{Rt}) + \\ &+ \int_{\pi_{M0}-\pi_{R0}}^{\pi_{M1}-\pi_{R1}} p_M(x_{i_t}, \pi_t) d(\pi_{Mt} - \pi_{Rt}), \end{aligned} \tag{9}$$

where

$$\begin{aligned} p_A(x_i, \pi_t) \equiv E[I_A(s_i, \pi_t)|x_i] = Pr[&-(u_{Ai} - u_{Ri}) < \pi_{At} - \pi_{Rt} + (\beta'_A - \beta'_R)x_i, \\ &-(u_{Ai} - u_{Mi}) < \pi_{At} - \pi_{Mt} + (\beta'_A - \beta'_M)x_i], \end{aligned} \tag{10}$$

is the propensity of an individual with observables x_i to work in the abstract task given the prevailing task price vector π_t (equivalently for $p_M(x_i, \pi_t)$). For Equation (9) to be true, the comparability assumption that the distribution of skills s_{i_1} and s_{i_0}, conditional on x_{i_1} and x_{i_0}, are the same has to hold. This is discussed in detail in the next subsection.

From Equation (9) I want to estimate the distances $\triangle(\pi_A - \pi_R)$, $\triangle(\pi_M - \pi_R)$, and $\triangle\pi_R$. $E(w_{it}|x_i)$, $p_A(x_i, \pi_t)$, and $p_M(x_i, \pi_t)$ are known at $t = 0$ and $t = 1$ in the sense that one can consistently estimate them from data with sufficiently detailed information about talents x_i (now t indexes two observed points in time in the data). However, $p_A(x_i, \pi_t)$ and $p_M(x_i, \pi_t)$ within the interval $t\epsilon(0,1)$ are unknown and one needs to make an assumption on them. I linearly interpolate

$$p_K(x_i, \pi_t) \approx p_K(x_i, \pi_0) + [p_K(x_i, \pi_1) - p_K(x_i, \pi_0)] \frac{(\pi_{Kt} - \pi_{Rt}) - (\pi_{K0} - \pi_{R0})}{(\pi_{K1} - \pi_{R1}) - (\pi_{K0} - \pi_{R0})} \tag{11}$$

with $K\epsilon\{A, M\}$, so that Equation (9) becomes

$$\begin{aligned} E(w_{i_1 1}|x_{i_1}) - E(w_{i_0 0}|x_{i_0}) = \triangle\pi_R + \frac{p_A(x_{i_t}, \pi_1) + p_A(x_{i_t}, \pi_0)}{2} \triangle(\pi_A - \pi_R) + \\ + \frac{p_M(x_{i_t}, \pi_1) + p_M(x_{i_t}, \pi_0)}{2} \triangle(\pi_M - \pi_R). \end{aligned} \tag{12}$$

For more details on the steps from Equation (7) to (12) via (9) see Appendix B.

Equation (12) is a new result in the Roy model in repeated cross-sections when only task prices are changing. The broader intuition behind the result is that for those workers who stay in their original task, the wage change over time that they experience is equal to the price change in this task. If workers switch, the total wage change for a given individual is the weighted sum of the price changes in all tasks that they spend time in, with the weights corresponding to the fraction of time spent in each task. The latter is what the propensity averages in Equation (12) represent for a whole set of workers with characteristics x_i.[11]

As long as there is a sufficiently strong relationship between observables and task choices ("first-stage", see below), the linear approximation (11) should be of second-order importance. This is because the initial and the final sorting provide tight economic bounds (i.e., $[p_K(x_i, \pi_0), p_K(x_i, \pi_1)]$) for the comparative advantage in task K of workers with observables x_i. The movement of $p_K(x_i, \pi_t)$ between $p_K(x_i, \pi_0)$ and $p_K(x_i, \pi_1)$ is only linearly interpolated within these bounds. The Monte Carlo simulations in Section 3.3 also show no approximation error in the considered settings.

The parameters in Equation (12) can be estimated in different ways. The easiest is to note that (12) is a linear conditional expectation function and that OLS regression

[11]One may think that this is complicated by the fact that at the time of switching the wage changes discretely because different tasks reward different skills differently. However, in the Roy model, at the exact instant of switching the worker reaches an indifference point (Envelope argument in Equation 6) and the wage in the origin and destination task are exactly the same. Hence, there is no discrete wage change due to the individual using different skills in different tasks. This is why the specific distribution of skills is not important here and the method is general in this respect. I would like to thank an anonymous referee for pointing this out.

provides the best estimator of linear conditional expectation functions (e.g., Angrist and Pischke, 2008). Pooling the data over two periods, and using time interactions for the regressors of interest, allows estimating the changes in conditional expectations over time. This is what one is after in Equation (12). So the equation to be estimated by OLS in becomes:

$$w_{i_t t} = \alpha_0 + \alpha_1 \overline{p}_A(x_{i_t}) + \alpha_2 \overline{p}_M(x_{i_t}) + \alpha_3 \times 1[t=1] + \\ + \alpha_4 \overline{p}_A(x_{i_t}) \times 1[t=1] + \alpha_5 \overline{p}_M(x_{i_t}) \times 1[t=1] + \varepsilon_{i_t t} \tag{13}$$

with $\overline{p}_K(x_{i_t}) \equiv \frac{p_K(x_{i_t},\pi_1)+p_K(x_{i_t},\pi_0)}{2}$. This "Propensity Regression" gives an estimate of the conditional expectations function of interest $\widehat{E[w_{i1} - w_{i0}|\overline{p}_A(x_i), \overline{p}_M(x_i)]} = \hat{\alpha}_3 + \hat{\alpha}_4 \overline{p}_A(x_i) + \hat{\alpha}_5 \overline{p}_M(x_i)$, since according to (12), $E(w_{i1}|x_i) - E(w_{i0}|x_i) = E[w_{i1} - w_{i0}|\overline{p}_A(x_i), \overline{p}_M(x_i)]$. Therefore, $\widehat{\triangle\pi_R} = \hat{\alpha}_3$, $\triangle(\widehat{\pi_A - \pi_R}) = \hat{\alpha}_4$, $\triangle(\widehat{\pi_M - \pi_R}) = \hat{\alpha}_5$. The Monte Carlo simulations in Section 3.3 confirm that this estimation approach works. It is convenient because one can use individual-level wages directly in the estimation, although one is interested in the change of a conditional expectation. This also facilitates the empirical test of Prediction 3, again using linear regression in Section 5.1 to estimate the wage growth of workers with different propensities to start in the abstract and manual tasks.

Notice that the OLS regression (13) has no structural interpretation in the model. It is only a conduit for estimating the sought-after conditional expectation function. Because it is not true from Equation (12) that $E(w_{it}|x_i) = E[w_{it}|\overline{p}_A(x_i), \overline{p}_M(x_i)]$ in levels, the parameters α_0, α_1, α_2, and ε_{it} have in fact no interpretation in terms of the RBTC-Roy model. If one insists on some interpretation of α_1, α_2, and the regression error ε_{it}, one could interpret them as partly reflecting worker's skill levels, as they are correlated with wage levels (again, note, the model makes no predictions about these wage levels).

One alternative approach to identifying the task prices is based on minimum distance estimation of Equation (12). Briefly, the idea is that Equation (12) holds for every element of the x_i vector, thereby delivering J different moment conditions when $x_i = [x_{1i}, ..., x_{ji}, ..., x_{Ji}]'$. Estimation then relates the moment conditions to each other in a quadratic form. The procedure is explained in detail in (Boehm, 2013), and its estimates are reported as an alternative in Table 5 below. While less straightforward than regression (13), one attractive feature of the minimum distance is that it also provides an over-identification test ("J-test") of the model, because Equation (8) will naturally not be exactly correct for every element x_{ji} in the data.

3.2 Identification Assumptions

Result (12) and the Propensity Regression (13) map the changing wage returns to x_i over time to the task choices that x_i is associated with. This approach takes the task model seriously because it is not worker characteristics that are priced in the labor market directly, but the tasks that these characteristics help carry out. I now discuss the assumptions about x_i that have to hold in order to correctly identify the task prices.

Empirically, the estimation approach for task prices proposed in Equation (13) requires data on worker talents x_i that fulfill the following two assumptions (remember sub-index i_t indicates an individual i observed in the data in t):

Assumption 1 (First-stage)**.** *The vector x_{i_t} predicts workers' task choices $p_K(x_{i_t}, \pi_t)$ in both periods of time $t \epsilon \{0, 1\}$.*

Assumption 2 (Comparability)**.** *Individuals with the same x_i vector are comparable over time. That is, for all x_{i_0}=x_{i_1}, the population distributions of unobservable skills $u_{Ki_0} = \delta_K' z_{i_0}$ and $u_{Ki_1} = \delta_K' z_{i_1}$, $K \epsilon \{A, R, M\}$, are the same.*

The purpose of the first-stage is to provide good predictors of task choice probabilities conditional on x_i and π_t. As their predicted values are usually similar, it is conceptually not important whether these probabilities are constructed using multinomial choice models, linear probabilities, or non-parametrically with cell means for each x_i realization.[12]

Economically, Assumption 1 however requires that for three tasks there exist at least two elements of the x_i vector that generate qualitatively independent task predictions. For example, in the NLSY data below, conditional on the other talents, the math talent predicts the abstract task (the element corresponding to math in β_A is high; compare Equation (10)) and the mechanical talent predicts the routine task (the element corresponding to mechanical in β_R is high). The manual (third) task can then be predicted by low math and low mechanical talent (in the data, in addition, high verbal talent together with low math talent predict the manual task).

The comparability Assumption 2 implies that the unobservable skill selection into observable groups x_i does not change over time. Thus, individuals with the same x_i vector are "in distribution" equally skilled in both periods ($F_{s_A, s_R, s_M | x_i}$ does not carry a

[12]In the NLSY data below, with several continuous worker traits, a multinomial logit model of occupational choice is fitted, while in the Census/ACS data, taken from Acemoglu and Autor (2011), the actual choice frequencies for discrete demographic x_i cells are used. Alternatively, a linear probability model would require two different choice regressions when there are three tasks and the predicted probabilities would not be bounded by 0 and 1.

time index). Assumption 2 also implies that $p_K(x_i, \pi_t)$ functions are time-invariant with respect to their x_i argument holding constant π_t. If one observes an individual with x_{i_1} in $t = 1$, his counterfactual employment in tasks in $t = 0$ would have been the employment of an individual with the same x_{i_0} in $t = 0$.

The NLSY dataset is attractive in terms of comparability because it provides pre-labor market talents x_i that are difficult to influence for an individual and which have hardly changed (in levels and correlations) over time (Altonji, Bharadwaj, and Lange, 2012; Speer, 2017). If the population distribution of unobservable talent is constant across cohorts, and there are no groups in terms of unobservables who behave sub-optimally and acquire less of a characteristic that becomes more desirable over time (i.e., decrease their math or verbal test scores), comparability should hold in that data.

The Monte Carlo simulations in the next section examine to what extent the Propensity Method can allow for confounding factors that affect wages. Broadly, some types of changing skill accumulation or changing returns to skills can be accounted for separately, but not when they occur together. Although this a potentially important limitation of the Propensity Method (it could be interpreted as a violation of the comparability assumption), one advantage of using the young workers in the NLSY data is that at least the change in accumulated skill via task-specific experience is unlikely to be very strong at age 27.[13]

In the alternative dataset summarized at the end of the paper, the Census/ACS, education-age-region demographic cells are used as components of the x_i vector. While more debatable, I argue that also in this case the correct task prices may potentially be identified. This is because previous literature (e.g., Autor, 2015) has shown that males' educational attainment was remarkably slow to respond to differential wage premia, which is reflected in descriptive statistics below. Employment trends across regions were also modest, and aging in the population was mostly driven by birth rates decades earlier and thus unlikely to be affected by RBTC.

To save on notation, I suppress the t sub-index in i_t from now on, as in the actual data workers are always observed at the time when also the prices that they face are measured (i.e., the t index on π_t or w_{it} should suffice for clarity).

[13] Another advantage of using young workers, according to Gottschalk, Green, and Sand (2015), is that their wages better capture contemporaneous task price movements compared to older workers, whose wages are more sticky (e.g., because of implicit contracting).

3.3 Monte Carlo Simulations

This section conducts Monte Carlo simulations in order to assess the ability of the Propensity Method to recover the correct task prices. It also summarizes the results from introducing rising returns to college, increasing college attainment, and a changing minimum wage as potential confounders (details in Appendix C). An-easy-to-use Stata dofile is posted on my website for interested readers to replicate the results and to try out alternative parametrizations (e.g., for the unobserved skills).[14]

I simulate data that resembles what is observed in the actual NLSY below. Workers possess a vector of observed talents x_i, including math, verbal, and mechanical talent, which are jointly normally distributed and positively correlated among each other. A linear combination of these talents map into the (log) observable skill component in abstract, routine, and manual tasks. In particular, math and verbal load highly on abstract, mechanical loads highly on routine, and verbal loads *relatively* highly on manual. The unobservable skills in Roy-type models are often taken as normally or extreme value distributed (e.g., Heckman and Sedlacek, 1985; Hsieh, Hurst, Jones, and Klenow, 2013). Therefore, I let (log) unobservable skill components be either distributed multivariate normal, again with a positive correlation among each other, or type 1 (Gumbel) extreme value. The latter is a special case of the type 2 (Frechet) extreme value distribution for which relative task prices can be estimated using multinomial logit regressions.[15]

The model has a period $t = 0$ (corresponding to the NLSY79) and a period $t = 1$ (NLSY97). Abstract and manual task prices are set to rise between these two periods by 40 and 50 log points, respectively, and routine task prices decline by 15 log points. Workers' potential wages in each task and period are the sum of the log task prices, observable skill components, and unobservable skill components as in Equation (8). Workers choose the task that offers the highest potential wage in each period. I simulate the model using a small sample with 4,000 individuals (2,000 in each period), to mimic the NLSY, and a

[14]Go to "Code for Monte Carlo Simulations" on https://sites.google.com/site/michaelboehm1/research.

[15]To be precise, the assumptions about skills are:

$$\begin{array}{l} s_A = 6x_{math} + 2x_{verb} + 0x_{mech} + u_A \\ s_R = 0x_{math} + 0x_{verb} + 4x_{mech} + u_R \\ s_M = 0x_{math} + 1x_{verb} + 0x_{mech} + u_M \end{array} \quad \text{with} \quad \begin{pmatrix} x_{math} \\ x_{verb} \\ x_{mech} \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}, \begin{matrix} 1 & .7 & .5 \\ .7 & 1 & .6 \\ .5 & .6 & 1 \end{matrix} \right)$$

$$\text{and} \quad \begin{pmatrix} u_A \\ u_R \\ u_M \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}, \begin{matrix} 1 & .5 & .5 \\ .5 & 1 & .5 \\ .5 & .5 & 1 \end{matrix} \right) \quad \text{or type 1 extreme value.}$$

Period 0 task prices are $\pi_{A0} = 1$, $\pi_{R0} = 3$, $\pi_{M0} = .5$ in order to match the high and low initial employment share of the routine and manual tasks, respectively.

larger sample with 40,000 individuals (20,000 in each period), to have enough statistical power in order to identify (the likely size of) any potential bias.

There are 100 iterations. In each iteration I estimate the first-stage in each period separately using a multinomial logit regression of task choice indicators on the math, verbal, and mechanical talents x_i. I then construct the average task choice propensities $\overline{p}_K(x_i) \equiv \frac{p_K(x_i,\pi_1)+p_K(x_i,\pi_0)}{2}$ with $K \epsilon \{A, M\}$ from the predicted values of the multinomial regressions. Finally, I regress workers' observed wages onto these propensities interacted with time (i.e., Propensity Regression 13).

Table 2: Monte Carlo Simulations for the Basic Model

	Propensities Method			Multinomial Logit	
	$\triangle\pi_R$	$\triangle(\pi_A - \pi_R)$	$\triangle(\pi_M - \pi_R)$	$\triangle(\pi_A - \pi_R)$	$\triangle(\pi_M - \pi_R)$
TRUE	-15.0	40.0	50.0	40.0	50.0
Extr. value; 4,000 individ.	-15.73	40.45	51.07	41.21	48.55
St. Error of Mean	(1.37)	(2.88)	(2.40)	(1.28)	(1.32)
Avg Std Error	*(1.73)*	*(27.09)*	*(37.65)*		
Extr. value; 40,000 ind.	-14.53	39.09	49.62	39.46	49.84
St. Error of Mean	(0.47)	(1.03)	(0.96)	(.42)	(.40)
Avg Std Error	*(5.49)*	*(8.57)*	*(11.89)*		
Multiv. norm; 4,000 ind.	-14.91	37.13	51.25	72.39	93.99
St. Error of Mean	(1.51)	(2.97)	(2.46)	(1.98)	(1.85)
Avg Std Error	*(16.78)*	*(26.02)*	*(34.91)*		
Multiv. norm; 40,000 ind.	-14.71	40.67	49.83	74.70	91.86
St. Error of Mean	(.50)	(1.06)	(.72)	(.65)	(.60)
Avg Std Error	*(5.30)*	*(8.23)*	*(10.99)*		

Notes: The table reports the mean estimated task prices from the Propensity Regression (samples with 4,000 and 40,000 individuals) under extreme value type 1 and normally distributed unobservables. The standard error of the mean estimate over 100 iterations (in parentheses) and the average estimated standard error (in parentheses and italics) are also shown. For parametrization details, refer to the text and footnotes.

Table 2 reports the results from this exercise. Under both the extreme value (upper part) and the multivariate normal distribution (lower part) of unobserved skills, the Propensity Method recovers the true task prices very well. The multinomial logit regression gets even marginally closer to the actual task prices if the true unobservable skills are distributed extreme value type 1. But it is off by orders of magnitude if true task prices are distributed multivariate normal. In previous versions of this paper I also showed the converse finding, that is, that assuming multivariate normal unobservables yields incorrect task prices when true unobservables are extreme value distributed. In the Propensity Method, the standard error of the mean in the simulations (directly below the point estimate) does not let one reject the null hypothesis that the estimator recovers the true prices exactly, even in the large sample. This statistically supports the economic argument from Section 3.1 that the linear interpolation of the adjustment path, which yields Equation (12), is not a problem.[16]

I have checked the robustness of these results using different loadings of the talent vector into observable skill components, different joint distributions of unobserved skills (e.g., type 2 extreme value and uniform), different levels and changes of task prices, and using multinomial probits and linear probability specifications in the first stage of the Propensity Regression. The simulated model also yields reasonable aggregate outcomes, such as polarizing employment, and, under these specific parameters and evolution of task prices, a polarization of the overall wage distribution.

The second part of the Monte Carlo simulations examines to what extent the Propensity Method is compromised when potentially confounding forces impact the wage distribution (detailed analysis in Appendix C). First, a rising return to college is examined. Appendix C shows that an augmented Propensity Regression (13) can account for this:

$$\begin{aligned} w_{it} = \alpha_0 + \alpha_1 \overline{p}_A(x_i) + \alpha_2 \overline{p}_M(x_i) + \alpha_3 \times 1[t=1] + \alpha_4 \overline{p}_A(x_i) \times 1[t=1] + \\ + \alpha_5 \overline{p}_M(x_i) \times 1[t=1] + \alpha_6 c_i + \alpha_7 c_i \times 1[t=1] + \varepsilon_{it}, \end{aligned} \tag{13'}$$

where α_7 allows for the changing returns to college c_i. The specification can also account for task-specific returns to college. Second, increasing college attainment with no direct return to it (i.e., the strict task model where skills are only priced via tasks) is also not problematic. But if rising returns to college and increasing attainment occur together, there is no straightforward specification that always delivers the correct task prices. What seems important is to run both specifications, including and not including college, in regressions (13) and (13') to see whether differences arise. Finally, Appendix C also shows that the Propensity Method can be applied to modified data in order to account for

[16]The standard deviations of the price estimates for the small sample with 4,000 individuals are quite large (with 100 iterations, $\sqrt{100}$ times the standard error of the mean), and they are larger than for the multinomial logit. However, for 40,000 individuals the estimation becomes reasonably precise already, and the average standard error (reported one line below in italics) is at the same order of magnitude. In fact, in an earlier version of the paper with 2,000 simulation iterations, both are very similar. This suggests that inference even in the small sample of the NLSY is not compromised.

The average standard errors here are directly from the second-stage regression (13) without bootstrapping the first and second stage together. In the actual NLSY data, I also find that these "naive" standard errors are generally very similar to the bootstrapped standard errors, so that the bootstrapping correction does not make a big difference in practice.

changes in the minimum wage affecting the wage distribution.

To conclude from these Monte Carlo simulations, the basic version of the model without confounders recovers the changing task prices correctly under multiple tasks and for different joint distributions of workers' skills. Reassuringly, the linear interpolation approximation (11) is quantitatively unimportant and the OLS Propensity Regression (13) achieves identification. Moreover, while this paper analyses a task model and task prices, the potential confounders of rising returns to college, increasing college attainment, and changes in the minimum wage can also partly be accounted for.

3.4 Comparison to Alternative Estimation Methods in the Literature

This section compares the new Propensity Method for estimating task prices to existing alternative approaches. I argue that it differs from most reduced-form as well as structural methods in that it exploits a fundamental relationship between workers' wage growth and their task choices that underlies the economics of the Roy model. The Propensity Regression is also relatively easy to implement and transparent in terms of which empirical moments it exploits.

The methods of estimating task prices can broadly be categorized into reduced-form and structural approaches. The former's strategy is to control as well as possible for observable or time-invariant characteristics and then to invoke the assumption that residual or time-varying characteristics are not related to workers' task choices. Cortes (2016), for example, uses panel data in order to control for workers' task-specific fixed effects and observables characteristics (especially experience) in a wage regression. He then assumes that conditional task switching is unrelated to changes in potential wages across tasks and identifies the change in prices via time-varying task intercepts. Therefore, his "exogeneity" assumption implies that there is no idiosyncratic skill accumulation or learning about abilities that affects potential wages and makes individuals switch tasks.

Firpo, Fortin, and Lemieux (2013) use a recentered influence function regression to decompose changing inequality into a wage structure (skill and task prices) effect and an effect based on the changing supply of skills and tasks. They invoke an "ignorability" assumption, which states that conditional on the observable measures (especially education and experience), the distribution of unobservable skills remains constant within tasks. This assumption is the same as when interpreting the returns coefficients from a standard Oaxaca-Blinder decomposition as task prices, and Firpo, Fortin, and Lemieux (2013) argue that therefore their estimates are likely to be a lower bound of the true changes in task prices.[17]

Depending on the setting, such reduced-form approaches may be more or less successful at getting close to the actual task prices. But they are conceptually not fully satisfactory, because their estimation (though not the theory in the papers) assumes that there are no economic motives for switching into rising tasks conditional on the control variables. This raises the question why net switching into rising tasks is observed in the data, also conditional on controls, if not for economic reasons. The new Propensity Method presented in this paper is instead based on an economic model of worker skill selection, observable and unobservable, into tasks and the wage growth associated with it. Implementing this is demanding on the data because it requires that the distribution of unobservables conditional on observable worker characteristics does not change over time (comparability assumption 2), but it does not impose the strong requirement that the distribution of unobservables is the same conditional on task choice.

The strategy of structural methods for estimating task prices is to model the population distribution of unobservable skills explicitly instead of assuming it away. Examples for this are Heckman and Sedlacek (1985) and the literature that built on it. These papers impose a specific distribution of workers' unobserved skills (often joint normality) and then set up a likelihood function to estimate the task prices of interest, but also all the other variables of the model including the distribution parameters. As illustrated in the Monte Carlo simulations, the validity of the estimates from such methods depends on the correctness of these distributional assumptions. In contrast, the Propensity Method proposed in this paper is derived under any arbitrary distribution of workers' skills, while it does require comparability.[18]

There also exist intermediate ("semi-structural") approaches for estimating task prices, which are closer in spirit to the new method described in this paper. First, instead of relying on the normality assumption about unobservables, one could propose an estimation approach based on instrumental variables or an identification-at-infinity argument in the Heckman selection equation (e.g., Dahl, 2002; Mulligan and Rubinstein, 2008, respectively). That this has not been tried in the current context reflects that it is difficult to find instruments which credibly affect task choices but not potential wages or to construct a

[17]Qualitatively consistent with this paper, Cortes finds that abstract task prices rose by about 30 percent and manual task prices rose by 15 percent compared to routine task prices. Firpo, Fortin, and Lemieux find that technology (i.e., RBTC) played a central role during the 1980s and 1990s, while offshorability became an important factor from the 1990s onwards.

[18]It also contains the linear interpolation (11), but there are strong economic reasons for this approximation to be of second order importance compared to the economic content of workers' task choices, which is confirmed in the Monte Carlo simulations.

plausible identification-at-infinity argument. The comparability assumption in this paper is arguably easier to fulfill. One limitation compared to the structural models is that the Propensity Method can only estimate the changing intercepts (tasks prices), while the β_K coefficients of how worker abilities map into tasks are assumed time-invariant.

Yamaguchi (2016) presents another approach that could be considered semi-structural. In his model occupations are characterized, and skills are priced, exclusively by their complexity in a multidimensional vector of K observable tasks (in practice $K = 2$, with cognitive and motor tasks). The strength of this approach is to reduce many discrete occupations into a finite-dimensional and economically interpretable task-space. Yamaguchi (2016) estimates the model using correlated random effects, which express workers' unobserved time-invariant skills as a function of all observed characteristics and their labor market history. This yields changes in task-specific intercepts, which may be interpreted as task prices, and in slopes, which may be interpreted relationships between skills and tasks. The identification relies on the assumption that unobserved skills can be expressed as a function of all observables plus an "exogeneity" assumption, as in Cortes (2016), that idiosyncratic skill shocks or learning about skills are unrelated to task choices.[19]

Finally, Gottschalk, Green, and Sand (2015) use a bounding exercise to purge changing selection from wages in task in order to identify the task prices. They get a first set of bounds for each task by noting that, in a hierarchical skill model, there will be only worker flows between two neighboring occupations in terms of skills (i.e., between abstract and routine or routine and manual, but not between abstract and manual). Statistically, the probability that movers' wages are above or below the occupation median is between zero and one. So a wide set of bounds tracks changes in the median wage under these two extremes. The authors tighten these bounds by appealing to the economic argument about stochastic dominance of skills between movers and stayers in tasks. They then relax the distribution of skills to alternative ability models and repeat the two steps. Gottschalk, Green, and Sand (2015)'s approach allows them to separately identify skill selection and changes in occupation-specific skill functions by appealing to higher moments of wages (i.e., additional percentiles on top of the median). But the correct identification of bounds ultimately relies on the imposed restrictions for workers'

[19]Yamaguchi finds that the returns to motor tasks have declined, hurting male compared to female workers, while returns to cognitive tasks have been unchanged.

In an earlier paper, Yamaguchi (2012) estimates a similar model with the Kalman filter, relying in the estimation on the assumption that occupations are indeed fully described by the two observable task dimensions (cognitive and motor), and on the functional form of the normal distribution for skills, skill shocks, preferences, and measurement error in order to compute the likelihood function.

skill distributions (the most general one in the paper is a combination model of hierarchical together with independent worker-specific skills across occupations).[20]

Aside from the specific assumptions about the skill distributions, the bounding approaches and the structural estimations become tedious or computationally demanding with three or more tasks. Therefore, another attractive feature of the method proposed in this paper is that it is easily applied to three or more tasks. One just needs to include the respective task propensities into Regression (13). The method is also transparent in terms of which moments in the data it uses for identification (see Equation 12).

4 Data

4.1 Occupational Grouping and Alternative Hypotheses

The new Propensity Method for estimating task prices requires discrete occupation groups. This has the advantage of being applicable to a broad set of problems (it is the classic Roy model), but also the cost that task input of every worker is not finely measured. Alternative hypotheses to RBTC, including trade/offshoring and SBTC coupled with rising demand for services, are also discussed in this section.

Acemoglu and Autor (2011) propose an occupational grouping based on two-digits of the 1990 Census Bureau occupational classification, which is consistent over several decades. Four groups are generated: (1) managerial, professional and technical occupations; (2) sales, clerical and administrative support occupations; (3) production, craft, repair, and operative occupations; (4) service occupations. Acemoglu and Autor then show that group (1) is intensive in non-routine cognitive, analytic, and interpersonal tasks according to the Dictionary of Occupational Titles (DOT) and the Occupational Information Network (O*NET), the two most-used sources for measuring occupational tasks. Groups (2) and (3) are intensive in routine cognitive and routine manual tasks, respectively. Group (4) is intensive in the non-routine manual task. In their empirical application, which is based on Prediction 3 above, Acemoglu and Autor further merge categories (2) and (3) into one, so that they end up with one abstract group (1), one routine group (2+3), and one manual/services group (4).

Acemoglu and Autor (2011)'s grouping into broad occupations carries limitations. First, the fine variation in tasks across more detailed occupations is aggregated away,

[20] Gottschalk, Green, and Sand find that task prices are polarizing until the year 2000, and that all task prices are falling thereafter.

and when one moves from four to three groups even the cognitive and manual routine groups are merged into one. Second, other studies have noted deviations within and changes from the routine task framework in these groups. Autor and Dorn (2013), for example, have found that transport, construction, mechanical, and mining are intensive in manual routine as well as non-routine tasks. They also argue that clerical and sales occupations are very diverse in their tasks and that clerical occupations' task content may have upgraded over time. Other studies construct their own task measures and groupings, for example based on a principal component analysis of the broader DOT or O*NET information (e.g., Firpo, Fortin, and Lemieux, 2013; Yamaguchi, 2012).

Recognizing that every occupational grouping involves some degree of subjectivity, I adopt the Acemoglu and Autor (2011) classification because it captures well the different intensities in abstract, routine, and manual tasks and because it has been used in many of the subsequent papers analyzing RBTC in the U.S. (e.g., Mishel, Shierholz, and Schmitt, 2013; Autor, 2015; Gottschalk, Green, and Sand, 2015; Cortes, 2016; Bárány and Siegel, 2017).[21] This enables a consistent comparison of the paper's findings with prior literature. It should be clear, however, that strictly speaking the task prices that are estimated are prices paid for an efficiency unit of labor in the three broad occupation groups, which in actuality constitute a bundle of all tasks at differing intensities. Another potential limitation, which is shared with virtually all papers in this literature, is that the task content even of detailed occupations may have changed over time.

The grouping of occupations is related to some alternative or complementary factors that may have affected task prices. First, there is a long-standing debate about the impact of international trade and offshoring on the wage and employment structure. Several authors have proposed measures of task content that may capture offshorability of occupations, either largely focusing on the importance of face-to-face contact and the need for on-site work in O*NET data (e.g., Firpo, Fortin, and Lemieux, 2013) or from directly surveying workers and occupational experts (Blinder and Krueger, 2013). These offshorability measures are mildly positively correlated with routineness and non-routineness in-

[21] As in Acemoglu and Autor (2011), occupations are first converted from their respective scheme into a time-consistent classification. They are then assembled into ten occupation groups, which are further aggregated into an abstract category (professional, managerial, and technician occupations); a routine category (sales, office/admin, production, and operator/labor occupations); and a manual category (protective, food/cleaning, and personal care occupations). *(Non-Routine) Manual*: housekeeping, cleaning, protective service, food prep and service, building, grounds cleaning, maintenance, personal appearance, recreation and hospitality, child care workers, personal care, service, healthcare support. *(Cognitive and Manual) Routine*: construction trades, extractive, machine operators, assemblers, inspectors, mechanics and repairers, precision production, transportation and material moving occupations, sales, administrative support. *(Non-Routine) Abstract*: managers, management related, professional specialty, technicians and related support.

dices (Firpo, Fortin, and Lemieux, 2013; Autor and Dorn, 2013). In the case of the broad Acemoglu and Autor (2011) occupations, the authors report that abstract and routine cognitive (1 and 2 above) are relatively offshorable, while routine and non-routine manual (3 and 4 above) are not. Autor and Dorn (2013) and Goos, Manning, and Salomons (2014) find that offshorability plays a minor role in explaining job polarization compared to routineness, while Firpo, Fortin, and Lemieux (2013) find that it is important for the wage structure from the 1990s onward.

Another hypothesis that has received attention is a combination of skill-biased technological change (SBTC) and rising demand for services. SBTC cannot explain the polarization of the employment or wage distribution by itself, since it predicts an increase in the relative demand for skill across-the-board (Acemoglu and Autor, 2011). But if SBTC raises incomes overall or at the top of the distribution, and consumption preferences are non-homothetic (e.g., high-income individuals substitute market for home-based production), this may raise demand for low-skill services occupations that are intensive non-routine manual tasks. Manning (2004) and Mazzolari and Ragusa (2013) present evidence that is consistent with this channel.[22] However, SBTC that works independent from RBTC would also imply a rising college premium (Acemoglu and Autor, 2011), and Section 5.1 in this paper finds very limited evidence for this conditional on task propensities.

In conclusion, the overall body of evidence suggests RBTC to be the most important driver of polarization, at least in the employment structure. While it seems hard to fully separate out the alternative explanations, the estimates in the following provide prices for occupation groups that employ every task to some extent but are intensive in abstract, routine, and manual, respectively. The previous section has also shown how certain confounding factors such as a rising college premium or changing minimum wage can be directly controlled for in the Propensity Regression. These alternatives, and also the hypothesis that skill supplies have driven the changes, do not receive strong empirical support below.[23] In what follows, I refer to the three occupation groups by *abstract*, *routine*, and *manual tasks*.

[22] Bárány and Siegel (2017) propose a related explanation in which structural transformation across sectors drives the rising demand for services occupations. Acemoglu and Autor (2011) and Goos, Manning, and Salomons (2014) show that jobs also strongly polarized within industry sectors.

[23] Firpo, Fortin, and Lemieux (2013) analyze changing union membership as another institutional factor and find it to be important during the 1980s and 1990s.

4.2 The NLSY Sample of 27 Year Old Males

This section explains why the analysis focuses on males, introduces the NLSY sample, and computes the main facts concerning the distributions of jobs and wages therein.

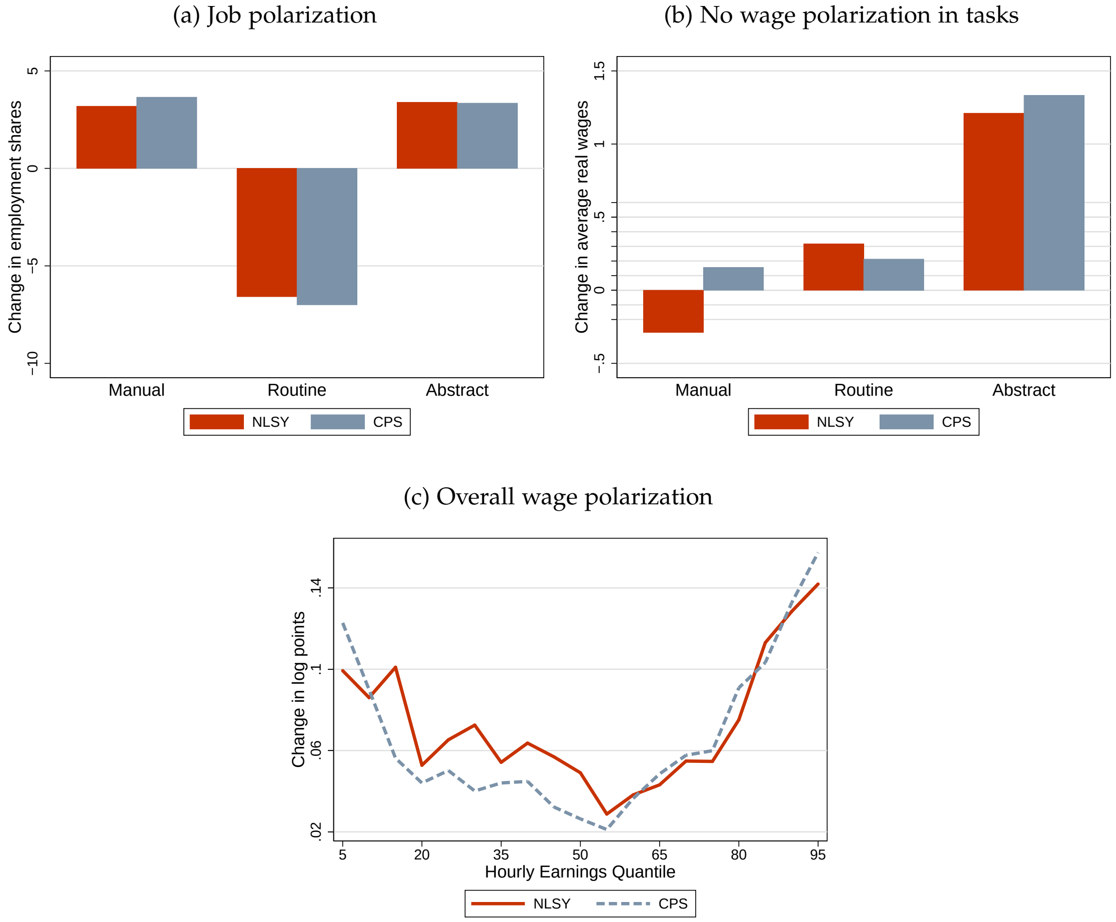
First of all, notice that the identification assumptions will be more plausibly fulfilled for male workers than for females. Female educational attainment as well as their participation in the labor market have much increased over the last decades. In fact, even for the different test scores, female performance improved noticeably between the two cohorts of the NLSY while male performance remained constant. Therefore, the comparability assumption is more likely to be violated with the available characteristics x_i for females than for males. Moreover, female wages rose substantially across-the-board compared to males and some argue that discrimination against them in different high-skill occupations has declined quite drastically (e.g., Hsieh, Hurst, Jones, and Klenow, 2013). Therefore, it is likely that a large part of the returns to characteristics x_i is driven by other factors than RBTC and task prices. Finally, the "mechanical" talent affects female task choices in the NLSY97, but not in the NLSY79. This is a problem for the first-stage regression (Assumption 1) and it suggests a violation of comparability (Assumption 2). For these reasons, the analysis in the main text is restricted to males. Appendix F summarizes estimates for females.

The main sample that I use is from the two cohorts of the National Longitudinal Survey of Youth (NLSY) and, for comparison, from the Current Population Survey Outgoing Rotation Groups (CPS) over the same period. I focus on 27 year old males in 1984–1992 and 2007–2009 in the NLSY 1979 and 1997, respectively. The details of the sample construction can be found in section D of the Appendix.[24]

Figure 2 presents the labor market facts of 27 year olds between 1984–1992 and 2007–2009 for the NLSY and the CPS. Employment polarized substantially during this period (Panel (a)). However, Panel (b) shows that average wages in the manual task hardly increased in the CPS and fell in the NLSY such that wages in tasks did not polarize. As

[24] The sample selection and attrition weighting for the NLSY data is done closely in line with a recent paper by Altonji, Bharadwaj, and Lange (2012). Since attrition in the NLSY97 is higher and test taking is lower than in the NLSY79, Altonji, Bharadwaj, and Lange (2012) examine it in detail. They conclude that after appropriate sample weighting any potential biases are not forbidding. I do not use the 2010 and 2011 samples of the NLSY97 because wages are substantially lower and less abstract (more manual) tasks are chosen compared to the CPS. Also, the AFQT scores of those members of the 1983–84 birth cohorts who work as 27 year olds in 2010–11 are substantially lower than the AFQT scores of the working 1980–82 birth cohorts. I construct labor supply by hours worked and real hourly wages as in Lemieux (2006). Table 7 in the Appendix accounts for how I end up with a sample of 3,054 and 1,207 individuals in the NLSY79 and the NLSY97, respectively.

Figure 2: The distributions of employment and wages for males age 27 in the NLSY and the CPS (1984/92 to 2007/09)



argued in relation to simulation Figure 1, this could be due to changing selection bias into the manual task even under RBTC and task price polarization. Finally, the overall wage distribution polarized substantially, both in the NLSY and in the CPS (Panel (c)).

4.3 Comparability and First-Stage in the NLSY

The attractiveness of the NLSY data for applying the Propensity Method is that it provides measures of workers' early skill determinants ("talents") that are not available in other datasets. These talents are determined pre-entry into the labor market and relatively hard to change for an individual since they are constructed from different components of an aptitude test. As elements of the x_i vector, the NLSY talents therefore come as close as possible to fulfilling the comparability Assumption 2.

I construct measures of mathematical, verbal, and mechanical talent by using test scores on mathematics knowledge, the average of paragraph comprehension and word knowledge, and the average of mechanical comprehension and auto- and shop information, respectively, from the components of the Armed Services Vocational Aptitude Battery of tests (ASVAB). A similar definition of talents has been adopted by a couple of subsequent papers on the education and labor market effects of different worker abilities in the NLSY (e.g., Prada and Urzúa, 2017; Speer, 2017).[25]

Table 11 in the Appendix presents labor force averages of talents as well as some demographic variables and contemporary skill measures that are available in more standard datasets. Following Altonji, Bharadwaj, and Lange (2012), the AFQT scores are adjusted for differences in test taking age and the switch from paper-based to computer-administrated tests between the NLSY79 and the NLSY97. In Table 11 the mean of this comparable-over-time proxy for general intelligence hardly changes between the two cohorts (standard deviation is 31.7 in the NLSY79 and 32.2 in the NLSY97).[26] In addition, Table 3 reports that the cross-correlation of the composite test scores and AFQT remained virtually the same. Taken together, the two tables show that the joint distribution of talents remained stable over time, which lends support to the comparability assumption 2 for them as components of the x_i vector.[27]

Figure 3 depicts average mathematical, verbal, and mechanical talent in the three tasks in both cohorts. The levels of the three talents are substantially higher in the abstract task than in the routine task which, in turn, is higher than the manual task. Thus, there is a clear ordering of absolute advantage in tasks independent of the talent considered. However, in the absence of restrictions to enter tasks, workers' choice should be

[25] All the measures used here are taken from the ASVAB, which consists of ten components: arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, general science, numerical operations, coding speed, auto and shop information, mechanical comprehension, and electronics information. The breakup into mathematical, verbal, and mechanical talent is similar to what a factor analysis of the test scores suggests. AFQT is essentially the average of arithmetic reasoning, word knowledge, paragraph comprehension, and mathematics knowledge.

[26] One early determined characteristic that is not constant is the share of Hispanics, which rose by 8 percentage points. I therefore control for race in all analyses. Also, when excluding Hispanics from the dataset, the cross-correlation between talents remains virtually the same in both cohorts while the level of AFQT rises from 168.5 to 169.7 over time. This is still very small given AFQT's standard deviation.

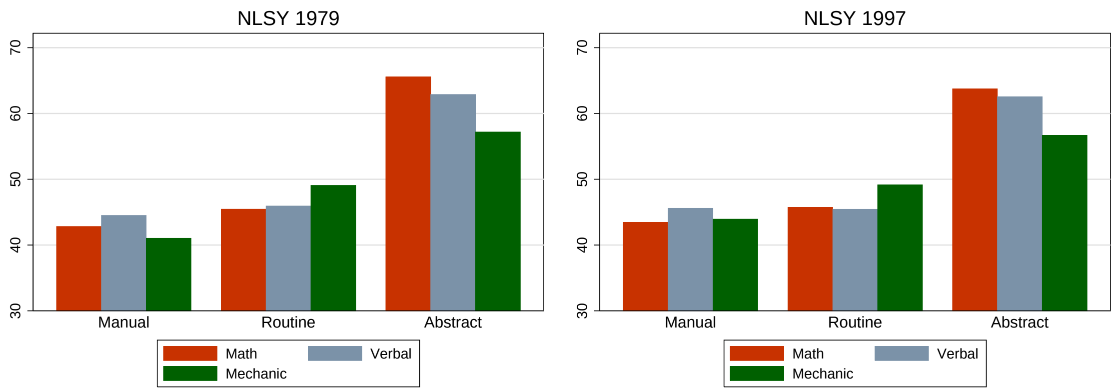
[27] In the later cohort of the NLSY, for which the tests are taken at age 12–16 in 1997, one concern may be that individuals endogenously invest into their talents as a response to RBTC. What would be required for a violation of the comparability assumption here is not that more able students generally achieve higher test scores, but that students increase their math and verbal scores, which predict abstract and manual tasks, in response to RBTC already before age 12–16. In this case, also some students need to behave sub-optimally and decrease their test scores, as the level of AFQT and the cross-correlations of talent measures is stable between the NLSY cohorts. While one may debate this possibility, it is also not clear whether high school students and their parents were even aware of the shifts in task demands that were going on by 1997 as, for example, the first academic papers about this phenomenon by Autor, Levy, and Murnane and Goos and Manning were only published in 2003 and 2007, respectively.

Table 3: Pairwise Correlations between Talents, NLSY 1979 and 1997

	NLSY79			NLSY97		
	AFQT	Math	Verbal	AFQT	Math	Verbal
AFQT (NCE)	1			1		
Math Score (NCE)	0.82	1		0.83	1	
Verbal Score (NCE)	0.93	0.71	1	0.92	0.75	1
Mechanical Score (NCE)	0.63	0.53	0.61	0.63	0.54	0.63
Nbr Observations	2,936			1,207		

Notes: The table shows the pairwise correlations between composite test scores after standardizing to normal curve equivalents with mean 50 and standard deviation 21.06.

Figure 3: Average talents in tasks, NLSY 1979 and 1997



governed by their comparative advantage and thus depend on their relative skills. This principle seems to be borne out in Figure 3. Average mathematical talent in the abstract task is higher than average verbal or mechanical talent, while average mechanical talent is considerably higher in the routine task than mathematical or verbal talent. Verbal talent is higher than mathematical and mechanical talent in the manual task.

Table 12 in the Appendix quantifies the sorting of talents into tasks. First, Columns (1) and (3) run multinomial logit regressions of task choice onto the linear talent measures, extracting the marginal effect of an additional unit of each talent on choosing the abstract and manual task relative to the omitted routine task. In Column (1), conditional on the other talents, a one unit higher math score is associated with an about 4.7 percentage point higher probability to enter the abstract versus the routine or the manual task. A one unit higher mechanical score is associated with a 1.4 and 2.3 percentage point lower probability to enter the abstract and the manual task as opposed to the routine task, respectively. In contrast, a one unit higher verbal score decreases the probability to enter the routine as opposed to the abstract or the manual task by about two percentage points. These results support Assumption 1 on the first-stage and they are similar in the NLSY97 in Column (3) of the table.

The first-stage estimates for the Propensity Method are from a more flexible version of this regression. First, I include terciles in math, verbal, and mechanical talent in order to allow for the fact that absolute advantage is partially aligned with relative advantage in the NLSY data (i.e., all three measures are high in abstract tasks in Figure 3). Normalized measures of illicit activities under age 18 and engagement in precocious sex are also added to capture differences in task choice according to non-cognitive traits (the popular locus of control and self-esteem are not available in the NLSY97). The estimates reported in Columns (2) and (4) of Table 12 show that the relationship for math, verbal, and mechanical talent is similar to above, with the limitation that not all tercile coefficients are significant at the five percent level. In addition, engagement in illicit activities during youth predicts not working in the abstract task conditional on the other talents.

5 Results

This Section first tests Prediction 3 in the NLSY data. Then it presents the Propensity Method's estimates of the changes in task prices. Finally, the effect that task prices may have had on the overall wage distribution is assessed.

5.1 Workers' Wage Growth in the NLSY over the 1990s and 2000s

Theoretical Prediction 3 states that relative wages of workers who start out in routine tasks will decline when task prices polarize. Since the same individual workers are not in the NLSY79 and NLSY97, Prediction 3 has to be evaluated using predicted probabilities based on talents x_i analogous to the Propensity Method for estimating task prices.

Focusing only on workers' initial task choices gives a reduced-form regression equation of the form:

$$\begin{aligned} w_{it} = \alpha_0 + \alpha_1 p_A(x_i, \pi_0) + \alpha_2 p_M(x_i, \pi_0) + \alpha_3 \times 1[t=1] + \alpha_4 p_A(x_i, \pi_0) \times \\ \times 1[t=1] + \alpha_5 p_M(x_i, \pi_0) \times 1[t=1] + \alpha_6 c_i + \alpha_7 c_i \times 1[t=1] + \varepsilon_{it} \end{aligned} \quad (14)$$

This regression differs from the Propensity Regression (13) for estimating task prices in that it uses only the period $t = 0$ task choice probabilities (i.e., fitted values for the NLSY79 and counterfactual predicted probabilities for the NLSY97 data). The Monte Carlo simulations with confounders (see Equation 13′ and Appendix C) have shown that a variable c_i and its time interaction can be included in the regression to control for changing returns to college that may affect workers' wages aside from RBTC. According to Prediction 3, at least one of the parameters α_4 and α_5 should be positive.

Analogous to the Propensity Method, the first-stage and comparability assumptions need to hold in order for the regression to identify these parameters correctly. The first-stage is estimated in a multinomial logit regression and presented in Column (2) of Table 12 (see interpretation in previous section). Multinomial probit or linear probability models give similar results. The predicted values, that is, the regressors in Equation (14), are remarkably stable over the two cohorts of the NLSY (reported only in earlier versions of this paper and available upon request). This supports the comparability assumption, as counterfactual labor supply into tasks according to observables remained largely unchanged. The first-stage multinomial choice regression and the second stage wage regression are bootstrapped in order to obtain the correct standard errors given that $p_A(x_i, \pi_0)$ and $p_M(x_i, \pi_0)$ are estimates with sampling variation.

Table 4 displays the results from the second stage regression on task propensities (14). Unsurprisingly, in column one a higher propensity to enter the abstract task compared to the omitted routine task is associated with a significantly higher wage. A ten percentage point higher probability to enter an abstract task (rather than a routine one) is associated with a 3.1 log points higher wage. The reverse is true for the propensity to enter the manual task. RBTC should however change the returns to propensities over time (i.e., α_4 and α_5), which are indicated in the table by "x NLSY97". Indeed the coefficients change strongly and significantly in the direction of Prediction 3. For the propensity to enter the abstract task, the coefficient almost doubles (from .31 to .60) while the coefficient for entering the manual task rises by more than a third (from -1.65 to $-.95$).[28]

An important competitor hypothesis to RBTC, discussed in Section 4.1, is a combination of SBTC together with rising consumption demand for services. Since SBTC implies a rising return to college (e.g., Acemoglu and Autor, 2011), the remainder of Table 4 examines theoretical Prediction 3 when changing returns to education are allowed for. First, column two inserts a four-year college dummy and its time interaction into the estima-

[28]Previous versions of this paper (e.g., Boehm, 2013) included a figure showing that for individuals with a high propensity to enter the routine task, which is quite frequent in the data, predicted real wages even decline during the two decades between the NLSY79 and the NLSY97.

Table 4: Returns to NLSY79 Task Propensities over Time

	(1) Log Wage	(2) Log Wage	(3) Log Wage	(4) Log Wage
Propensity Abstract Task	0.31	0.03		-0.06
	(0.07)	(0.08)		(0.08)
Propensity Abstract Task x NLSY97	0.29	0.25		0.30
	(0.11)	(0.13)		(0.13)
Propensity Manual Task	-1.65	-1.80		-1.75
	(0.17)	(0.17)		(0.17)
Propensity Manual Task x NLSY97	0.70	0.86		0.91
	(0.39)	(0.38)		(0.38)
College		19.23	24.88	
		(2.92)	(2.45)	
College x NLSY97		4.04	8.44	
		(5.20)	(4.14)	
Observations	4154	4149	4260	4149
R^2	0.09	0.11	0.07	0.12
Degree Dummies x NLSY97	No	No	No	Yes

Notes: The table reports OLS wage regressions of 100 times the deflated log wage on predicted propensities to the abstract and manual tasks. The propensities are estimated from the NLSY79 only according to Column (2) in Table 12. "x NLSY97" stands for the change in the coefficient between the NLSY79 and the NLSY97. Bootstrapped standard errors (500 iterations) below the coefficients.

tion ($\alpha_6 c_i + \alpha_7 c_i \times 1[t = 1]$ in Equation 14). On the one hand, the level of the coefficient on the propensity to enter the abstract task drops all the way to zero, but the changes in both coefficients are remarkably stable. On the other hand, the level of the return to college is large and highly significant, while its change is not significantly positive once the propensities are accounted for (note that in column three the college dummy rises significantly).[29] The result is similar in the last column, which controls for four different degree dummies (high school dropout and graduate, some college, and at least four-year college) and their time interactions.

These results suggests that Mincerian returns to education are important to explain wages in the cross-section, but that they (and SBTC) seem to have less power than relative skills in tasks to explain the change in wages that took place over the twenty years from the NLSY79 to the NLSY97. Prediction 3 from RBTC is therefore supported by the results in Table 4, and rising returns to college may partly be a consequence of rising returns to abstract tasks.

[29]The increase in the college premium may even be overstated because of the net switching of workers into high college return tasks over time (see Appendix C).

Table 5: Estimated Task Price Changes in the NLSY (1984/92 to 2007/09)

	$\triangle(\pi_A - \pi_R)$ in log points (s.e.)	$\triangle(\pi_M - \pi_R)$ in log points (s.e.)	$\triangle\pi_R$ in log points (s.e.)	**Model Test (p-value in %)**
OLS on Propensities	25.1	32.9	-4.2	
	(12.3)	(38.4)	(7.7)	
OLS on Propensities	19.5	46.6	-6.4	
(2nd-stage college)	(13.7)	(37.9)	(7.8)	
OLS on Propensities	25.6	56.9	-5.2	
(2nd-st degree dum.)	(13.9)	(38.7)	(8.8)	
Opt. Min. Distance	20.2	38.9	-3.2	12.3
	(6.6)	(26.4)	(3.3)	(13.8)
OLS on Propensities	27.3	32.0	-5.7	
(Adj. for min. wage)	(12.7)	(40.0)	(8.0)	

Notes: The first row of the table presents estimated task price changes from the basic (13). The second and third row run the augmented Propensity Regression (13'), adding college and degree dummies interacted with time respectively. Row four reports the task price changes from a minimum distance estimation explained in detail in Boehm (2013). This also provides a test of the restrictions on talent returns implied by comparability and the RBTC-Roy model. The last row reports estimates when wages are first adjusted for the change in the real value of the minimum wage as in Lee (1999). Bootstrapped standard errors (500 iterations) below the coefficients.

5.2 Task Price Estimates

This section estimates the task price changes using different specifications in the NLSY data. Consistent with RBTC, task prices polarize over the joint 1990s and 2000s.

The first row of Table 5 reports the task price changes estimated from the basic Propensity Regression (13). As discussed above, the first-stage is estimated in a multinomial logit model with results reported in Appendix Table 12. For each individual, the averages of the predicted values in the period they are observed (actual) and the period when they are not observed (counterfactual) are used as the regressors. The whole procedure is again bootstrapped to obtain the correct standard errors for the second-stage estimates. The equilibrium prices being paid for tasks have changed substantially between the two NLSYs. In particular, the relative prices for the abstract and manual task increased by 25 and 33 log points, respectively, while the absolute price of the routine task decreased by 4 log points.[30] Task prices therefore polarized between the two cohorts of the NLSY with the qualification that the estimate for the manual task price is insignificant.

[30]Interpreting the intercept change in (13) as the absolute task price growth subsumes all time effects (e.g., also growth in aggregate productivity) under the task prices.

The remainder of Table 5 examines the robustness of this result. First, the augmented Propensity Regression (13′) is estimated, adding in rows two and three controls for college and degree dummies interacted with time as in Columns (2) and (4) of Table 4. The task price changes remain qualitatively the same in these specifications, with the relative price of the manual task somewhat increasing. The (unreported) rise of the college premium is insignificant as in Column (2) of Table 4 above. The task price polarization result therefore persists when allowing for rising returns to education implied by SBTC.

In row four of Table 5 the task price changes from an optimal minimum distance (OMD) estimation are reported. The OMD is introduced as an alternative approach to Regression for implementing the Propensity Method in Section 3.1 and it is explained in detail in an earlier version of this paper (Boehm, 2013). While less straightforward to implement, an advantage of the OMD estimate is that it does not suffer from potentially incorrect standard errors or attenuation bias due to sampling variation in the first-stage estimates $\hat{p}_K(x_i, \pi_t)$.[31] The (asymptotic) standard errors are lower due to its optimality. The OMD also provides an overidentifying restrictions test ("J-test") of the moment conditions implied by the empirical model, which is reported in the last column of row four. Reassuringly, the J-test does not reject the model and the point estimates in row four of Table 6 are similar to those in the previous rows, while the estimate for the change in the manual task price is now close to significant at the ten percent level.

Finally, one could be concerned about another force that might have worked aside from RBTC and confounded the task price estimates: the increase in the real value of the minimum wage in the U.S. between the end of the 1980s and the end of the 2000s. This may have raised wages in the lower end of the distribution as depicted in the bottom panel of Figure 2 and, since the manual task workers are more frequently found in this lower end, it may distort the task price estimates. The Monte Carlo simulations have shown that this confounder can be accounted for when one is able to restore the latent wage distribution that would have prevailed without the change in the minimum wage, even if individual workers do not get assigned their true latent wages.

Following Lee (1999), I therefore construct adjusted wages that would have prevailed in the absence of a change in the real minimum wage.[32] The wage distribution for 27 year

[31] The estimates are different from the ones reported in Boehm (2013), as an intercept was included to be consistent with the other specifications of Table 5. Attenuation bias due to sampling variation in the regressors is also not detectable in the Monte Carlo simulations.

[32] In order to generate wages in the face of a minimum wage at its 1989 level, I apply the method to compute the counterfactual from Lee (1999) together with the estimates of the effect of the minimum wage therein. Define the deflated minimum wage $\widetilde{mw}_t = (minwage_t - \bar{w}_t^t)$ as the minimum wage in t adjusted

olds in the NLSY and the CPS is now substantially flatter in the bottom than without the adjustment (compare the solid lines in Panels c and d of Figure 4 below). Row five of Table 5 presents the results from the task price estimation with the minimum wage adjustment. The price estimates remain similar to the preceding rows.

Overall, the findings in this section indicate that task prices polarized between the times when the members of the NLSY79 and the NLSY97 were 27 years old. In fact, allowing for changing returns to college and for a changing minimum wage does not qualitatively alter the estimates. The coexistence between polarizing task prices and polarizing employment also rules out the inverse hypothesis that task prices may have been driven by changes in relative labor supplies instead of labor demand. This points to the importance of RBTC in affecting workers' wages over this time period.

5.3 The Task Prices' Effect on the Overall Wage Distribution

One of the most debated questions in the literature on inequality is to what extent the demand for skills and tasks, the supply of skills, and policy factors have been responsible for the polarization of the U.S. wage distribution over the last couple of decades. This last section brings the wage analysis of RBTC back to the aggregate level, finding that the polarizing task prices can explain most of the increase of inequality in the upper half of the wage distribution and that they may have generated a flattening of the lower half. Minimum wages seem to have played an additional role in compressing the lower half of the wage distribution for younger workers in the U.S..

I assess the potential effect of task prices on the overall wage distribution by assigning every worker the price estimate for his task in the NLSY79:

$$\widehat{w_{i1}^{TP}} = w_{i0} + \widehat{\Delta\pi_R} + I_{Ai0}\Delta(\widehat{\pi_A - \pi_R}) + I_{Mi0}\Delta(\widehat{\pi_M - \pi_R}) \tag{15}$$

The predicted wage $\widehat{w_{i1}^{TP}}$ captures the effect of the task prices only. Within the RBTC-Roy model, the other factors that may affect overall wage inequality are shifts in skill

by the trimmed mean wage in the population where the bottom and top 30 percent of wages are removed (everything in logs). Then, analogous to Equation (9) in Lee, the amount

$$\triangle_{p,t} = \hat{\beta}_p(\widetilde{mw}_{1989} - \widetilde{mw}_t) + \hat{\gamma}_p(\widetilde{mw}^2_{1989} - \widetilde{mw}^2_t)$$

is added to a worker's wage in time t, where p denotes the worker's wage percentile, and $\hat{\beta}_p, \hat{\gamma}_p$ the estimated coefficients for the effect on each quantile reported in Lee's Table 1, Panel A, Column (5). Coefficients for the percentiles below the 10th and between the 10th and the 50th are linearly imputed. From the 50th percentile upward wages remain unadjusted as in Lee's paper. For example in the NLSY data, $\widetilde{mw}_{1989} = log(3.35) - 2.103$ where 3.35 is the nominal minimum wage in 1989 and 2.103 the trimmed mean log wage among 27 year old males in that year.

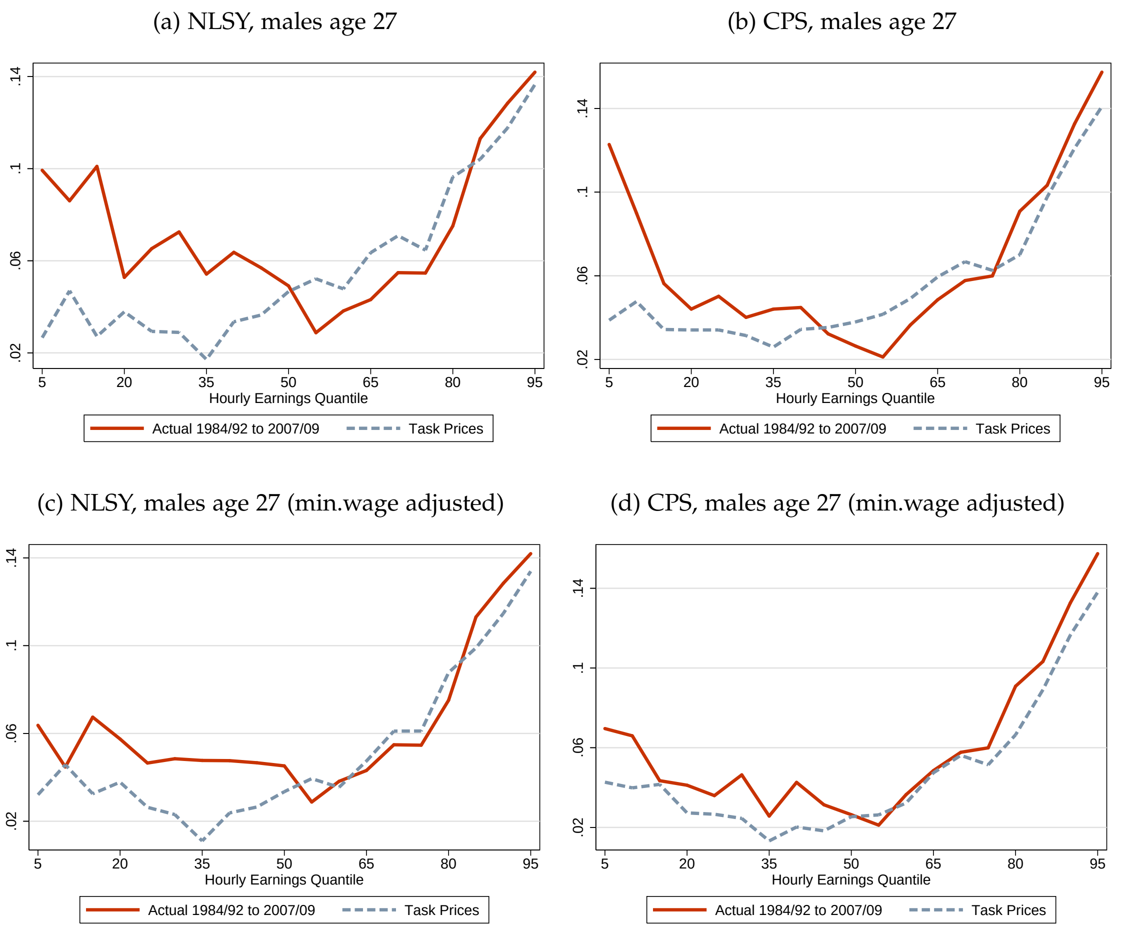
endowments and the wage effects of workers' task switching in response to the task price changes (to assess the latter effect, the population distribution of skills would have to be known). Outside the RBTC-Roy model, factors that could have affected wages include SBTC and policy or institutional variables such as changes in the minimum wage. The results so far suggest that SBTC and changing skill supply are not too important in explaining workers' wage growth conditional on skills in tasks, and that adjustments for the minimum wage do not affect the task price estimates. Nonetheless, it is unclear ex ante whether the task prices by themselves account for any substantial portion of the evolution of U.S. wage inequality.

The top row of Figure 4 plots the predicted change in the wage distribution (the quantiles of the distribution of $\widehat{w_{i1}^{TP}}$ minus the respective quantiles of the distribution of w_{i0}) together with the actual change in the wage distribution ("w_{i1} minus w_{i0}") for 27 year olds in the NLSY and CPS. The task price changes used in the predicted wage distribution, also in the CPS, are from the first specification in Table 5 (the specifications in rows 2–4 yield similar results). The prediction matches well the rise of the actual wage distribution in its upper half. However, the prediction is flat in the lower half and it does not account for the compression of inequality in that part of the actual wage distribution.

There are two reasons why the lower half of the overall wage distribution does not polarize from the changing task prices alone. First, Appendix Figure 6 plots the share of manual, routine, and abstract task workers into the NLSY79 wage distribution. It shows that routine workers are initially already concentrated in the lower half of the wage distribution, peaking around the 25th percentile (manual task workers are most concentrated in the very bottom). Therefore, declining routine task prices will strongest hit workers who already started out in the middle of the lower half of the wage distribution.

The second reason is the overtaking effect that was already discussed in relation to theoretical Prediction 2: manual task workers, who are predominantly located at the bottom of the wage distribution, move up under the new task prices. This lifts not only the low quantiles where the manual task workers start out, but also the more lower-middling quantiles of the wage distribution where they end up (i.e., routine task workers' initial position). The inverse happens for workers in routine tasks with the same effect on the wage distribution. This effect only exists in a truly multidimensional skill model. Appendix Figure 7 illustrates it, by plotting the predicted wage distribution when workers are fixed at their original quantiles so that overtaking is shut down. The increase is now weaker at the top and stronger at the bottom, since overtaking compounds the increase

Figure 4: Change in log real wages by quantile of the wage distribution, actual and predicted due to changing task prices (bottom row adjusted for the changing minimum wage)



of wages in the upper half and weakens the increase of wages in the lower half of the distribution when task prices polarize.[33]

As discussed in Section 2.2, the result that polarizing task prices do not lead to a clear polarization of the wage distribution is not necessarily evidence against RBTC. The overtaking effect illustrated in Appendix Figure 7 makes polarizing task prices increase inequality in the upper half of the wage distribution, while they only flatten (or even make more unequal, depending on the parameters) the lower half. Indeed, in most developed countries other than the U.S., the lower half or the wage distribution has not polarized and often it has become even more unequal during the last decades (see references in Section 2.2). The polarizing task prices and the changing wages of workers

[33]Note that overtaking not only exists when workers keep their original tasks (as in this section), but that it may also be substantial when one allows for the wage effects of switching tasks (as in Figure 1 Panel (d)).

who start out in different tasks, as estimated in the previous two sections, are therefore a better test of RBTC than their effect on the overall wage distribution.

Finally, one other factor that could have generated the distinct downward slope of the U.S. wage distribution, especially for the relatively young workers in the NLSY, is the increase of the minimum wage. The bottom row of Figure 4 plots the actual and the predicted distribution when wages are adjusted for the change in the real value of the minimum wage as in Section 5.2 and the task price estimates are taken from the corresponding specification in Table 5. The fit in the bottom of the wage distribution for 27 year olds is now substantially better. In the CPS, apart from a small difference in levels, the modest polarization in the lower as well as the increase in the upper half of the predicted and the actual wage distribution are now more or less comparable. In the NLSY, the difference is substantially reduced.

6 Robustness: the Census/ACS as an Alternative Sample

This Section briefly summarizes the results from an alternative estimation sample based on the decennial censuses and the American Community Survey (Census/ACS), which are broadly consistent with the findings in Section 5. The detailed analysis is relegated to Appendix E.

The strength of the Census/ACS is a large sample size and that the task price estimation can be done for different periods (decades). It also includes workers of broad ages 16–64. However, one has to to use education-age-region demographic cells as x_is for identification instead of the early-determined and time-invariant talent measures in the NLSY. This may be more of a concern for comparability Assumption 2.

Appendix Table 9 reports that also in the Census/ACS task prices polarized during the joint period 1989–2007. The effects are moderately stronger than in the NLSY and they are robust to including education, age, and region main effects as controls into the Propensity Regression, identifying task prices only from the interaction terms, and to adjusting for the minimum wage. The estimations by decade show that the relative price of the abstract task increased and the relative price of the routine task decreased during 1989–1999 and 1999–2007. The price of the manual task increased during 1989–1999, while its evolution over 1999–2007 is somewhat ambiguous. The task price estimates for the 1980s are rather inconclusive. Therefore, the clearest evidence for a polarizing impact of RBTC on workers' wages is during the 1990s.

On the aggregate level, the task prices match the changing overall wage distribution quite closely in the Census/ACS data. Appendix Figure 5 shows that only in the very bottom is the prediction not fully tracking the uptick of the actual wage distribution in the joint period 1989–2007. By decade, the polarizing wage distribution during 1989–1999 and the flat lower-half with rising upper-half of 1999–2007 are captured well, while the across-the-board increase in inequality during the 1980s is only partly accounted for.

Finally, Acemoglu and Autor (2011) estimate the wage growth of starters in abstract and manual tasks versus routine tasks in the Census/ACS. Similar to Section 5.1, they find that wages of the former types of workers rise compared to the latter over time, also when separately controlling for changes in the return to education (i.e., SBTC). Overall, therefore, while comparability in the Census/ACS is more debatable, these consistent results in another dataset strengthen the conclusions drawn from analysis of the NLSY.

7 Conclusion

This paper showed that in a Roy model changing task prices have robust implications for wage growth on the worker level, whereas changes in occupational or overall wage distributions are difficult to interpret. Exploiting this insight, a new Propensity Method for estimating task prices was derived.

Empirically, the paper finds that, consistent with RBTC, task prices polarized in the U.S. during the joint 1990s and 2000s. Workers with a relative advantage in routine tasks saw their wages decline compared to workers with a relative advantage in abstract and manual tasks. Notably, these are "ex ante" effects that are predicted by talents measured in high-school instead of actual experience in routine or non-routine tasks. Compared to existing studies analyzing the effect of technology or international trade using panel data (Cortes, 2016; Autor, Dorn, Hanson, and Song, 2014), this points to longer-run impacts of such shocks beyond the current generation of workers. The only other studies that have examined worker-level wages in the context of RBTC, and found broadly consistent results, are Cortes (2016) and Acemoglu and Autor (2011).

The task prices were then used to bring RBTC's effect on wages back to the aggregate wage distribution. This is therefore the only paper in the literature that establishes an explicit link between the actual predictions of RBTC theory and overall changes in wage inequality. Theoretically, polarizing task prices do imply that upper-part inequality should rise, but the effect on lower-part inequality is indeterminate. The empirical results suggest that the task price changes have led to a widening of inequality in the upper half of U.S. males' wage distribution and to a flattening of the lower half, but appear unable to explain all of the increase that is observed in the bottom. This is consistent with empirical evidence from other countries, where lower-half inequality did not narrow during the same time period.

More generally, the new theoretical results from the Roy model and the Propensity Method to estimate task prices are applicable beyond RBTC. For example, one may be interested in rising international trade competition affecting different occupations and industries (e.g., Autor, Dorn, Hanson, and Song, 2014), structural transformation across sectors (Young, 2014), or the evolution of employment demand for specific industries (Philippon and Reshef, 2012). Task prices are a very desirable quantity to obtain in such contexts, since they can be employed to compute labor supply schedules, to disentangle price and composition effects, to analyze effects on the overall wage distribution, and because they are a major empirical implication in their own right. In light of these different applications, an avenue of future research will be to extend the Propensity Method for use in longitudinal data, so that individuals' past task affiliations assume the role of talents in the estimation.

References

Acemoglu, D., and D. Autor (2011): "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings," vol. 4, Part B of *Handbook of Labor Economics*, pp. 1043 – 1171. Elsevier.

Adermon, A., and M. Gustavsson (2015): "Job Polarization and Task-Biased Technological Change: Evidence from Sweden, 1975–2005," *The Scandinavian Journal of Economics*, 117(3), 878–917.

Altonji, J. G., P. Bharadwaj, and F. Lange (2012): "Changes in the Characteristics of American Youth: Implications for Adult Outcomes," *Journal of Labor Economics*, 30(4), pp. 783–828.

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

Aughinbaugh, A., and R. M. Gardecki (2007): "Attrition in the National Longitudinal Survey of Youth 1997," Mimeo.

Autor, D. H. (2015): "Polanyi's Paradox and the Shape of Employment Growth," *Federal Reserve Bank of St. Louis: Economic Policy Proceedings, Reevaluating Labor Market Dynamics.*, pp. 129–177.

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

Autor, D. H., D. Dorn, G. H. Hanson, and J. Song (2014): "Trade Adjustment: Worker-Level Evidence," *The Quarterly Journal of Economics*, 129(4), 1799–1860.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006): "The Polarization of the U.S. Labor Market," *American Economic Review*, 96(2), 189–194.

——— (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90(2), 300–323.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118(4), 1279–1333.

Bárány, Z., and C. Siegel (2017): "Job Polarization and Structural Change," *Forthcoming at American Economic Journal: Macroeconomics*.

Blinder, A. S., and A. B. Krueger (2013): "Alternative Measures of Offshorability: A Survey Approach," *Journal of Labor Economics*, 31(2 Part 2), S97–S128.

Boehm, M. J. (2013): "Has Job Polarization Squeezed the Middle Class? Evidence from the Allocation of Talents," *Centre for Economic Performance Discussion Paper*, 1215.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality*," *The Quarterly Journal of Economics*, 128(3), 967–1015.

Cortes, G. M. (2016): "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data," *Journal of Labor Economics*, 34(1), 63–105.

Dahl, G. B. (2002): "Mobility and the return to education: Testing a Roy model with multiple markets," *Econometrica*, 70(6), 2367–2420.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009): "Revisiting the German wage structure," *The Quarterly Journal of Economics*, 124(2), 843–881.

Feenstra, R. C., and G. H. Hanson (1999): "The impact of outsourcing and high-technology capital on wages: estimates for the United States, 1979–1990," *The Quarterly Journal of Economics*, 114(3), 907–940.

Firpo, S., N. M. Fortin, and T. Lemieux (2013): "Occupational Tasks and Changes in the Wage Structure," *Working Paper*.

Goos, M., and A. Manning (2007): "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *The Review of Economics and Statistics*, 89(1), 118–133.

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

——— (2014): "Explaining job polarization: Routine-biased technological change and offshoring," *The American Economic Review*, 104(8), 2509–2526.

Gottschalk, P., D. A. Green, and B. M. Sand (2015): "Taking Selection to Task: Bounds on Trends in Occupational Task Prices for the U.S., 1984-2013," Working paper.

Green, D. A., and B. M. Sand (2015): "Has the Canadian labour market polarized?," *Canadian Journal of Economics/Revue canadienne d'économique*, 48(2), 612–646.

Heckman, J. J., and B. E. Honoré (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58(5), pp. 1121–1149.

Heckman, J. J., and G. Sedlacek (1985): "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, 93(6), pp. 1077–1125.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2013): "The Allocation of Talent and U.S. Economic Growth," Working Paper 18693, NBER.

Lee, D. S. (1999): "Wage Inequality in the United States during the 1980s: Rising Dispersion or Falling Minimum Wage?," *The Quarterly Journal of Economics*, 114(3), pp. 977–1023.

Lemieux, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 96(3), 461–498.

Manning, A. (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, F., and G. Ragusa (2013): "Spillovers from high-skill consumption to low-skill labor markets," *Review of Economics and Statistics*, 95(1), 74–86.

Mishel, L., H. Shierholz, and J. Schmitt (2013): "Don't Blame the Robots: Assessing the Job Polarization Explanation of Growing Wage Inequality," *EPI Working Paper*.

Moore, W., S. Pedlow, P. Krishnamurty, and K. Wolter (2000): "National Longitudinal Survey of Youth 1997: Technical Sampling Report," Discussion paper.

Mulligan, C. B., and Y. Rubinstein (2008): "Selection, Investment, and Women's Relative Wages Over Time," *The Quarterly Journal of Economics*, 123(3), 1061–1110.

Naticchioni, P., G. Ragusa, and R. Massari (2014): "Unconditional and Conditional Wage Polarization in Europe," *IZA Discussion Paper*.

Philippon, T., and A. Reshef (2012): "Wages and Human Capital in the U.S. Finance Industry: 1909-2006*," *The Quarterly Journal of Economics*.

Prada, M. F., and S. S. Urzúa (2017): "One Size does not Fit All: Multiple Dimensions of Ability, College Attendance and Wages," *Forthcoming at Journal of Labor Economics*.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

Speer, J. D. (2017): "Pre-market skills, occupational choice, and career progression," *Journal of Human Resources*, 52(1), 187–246.

Yamaguchi, S. (2012): "Tasks and Heterogeneous Human Capital," *Journal of Labor Economics*, 30(1), pp. 1–53.

——— (2016): "Changes in returns to task-specific skills and gender wage gap," *Journal of Human Resources*, pp. 1214–6813R2.

Young, A. (2014): "Structural Transformation, the Mismeasurement of Productivity Growth, and the Cost Disease of Services," *American Economic Review*, 104(11), 3635–67.

Appendix

A Proofs of Predictions 1-3

Sketch of Proof of Prediction 1. The change in average wages in task K can be split into a price and a selection effect:

$$E(w_i|I_K(s_i,\pi_1)=1)-E(w_i|I_K(s_i,\pi_0)=1)=\pi_{K1}-\pi_{K0}+E(s_{Ki}|I_{Ki1}=1)-E(s_{Ki}|I_{Ki0}=1)$$

While the (relative) prices $\pi_{K1}-\pi_{K0}$ may rise, the (relative) skills s_{Ki} selected into K may fall, depending on the overall distribution of worker skills in tasks. This is the classic idea of (changing) selection bias. In some cases overall change will be the inverse of the task price change. Figure 1 Panel (b) gives such a case and a counterexample that there need not be wage polarization in tasks (param. in Table 1). □

Sketch of Proof of Prediction 2. Focus on the lower half of the wage distribution. Consider manual task worker m and routine task worker r who are initially located at the 10th and 50th percentile of the wage distribution. For simplicity assume they do not switch tasks. If they stay at their original quantiles, the relative change in the quantiles becomes $\Delta w^{10}-\Delta w^{50}=\triangle\pi_M-\triangle\pi_R>0$, that is, we observe wage polarization. However, suppose the manual worker overtakes the routine worker (he benefits from the higher price change for the manual task) and that they exchange positions in the wage distribution. In this case the relative change in quantiles becomes $\Delta w^{10}-\Delta w^{50}=\underbrace{(w_{r1}-w_{m1})}_{-}+\underbrace{(w_{r0}-w_{m0})}_{+}$, which is flatter and may even be negative. Figure 1 Panel (d) provides a counterexample where the overall wage distribution does not polarize (parameters in Table 1). □

Sketch of Proof of Prediction 3. Consider Equation (7) and suppose that $\triangle\pi_A\geq\triangle\pi_M>\triangle\pi_R$. Every worker who starts in the A task will stay there and gain $\triangle\pi_A$. The stayers in the routine task will gain the smaller $\triangle\pi_R$. Even if they switch to the abstract or the manual task, none of the routine workers will gain the full $\triangle\pi_A$. These switchers will rather gain a weighted average of $\triangle\pi_A$ (or $\triangle\pi_M$) and $\triangle\pi_R$ (with nonzero weights depending on how quickly they switch), which is strictly smaller than $\triangle\pi_A$. However, it is theoretically possible that R starters (even on average) have higher gains than *either A or M* starters. For example, assume that $\triangle\pi_A>\triangle\pi_M$ and that (some of) the initial manual workers can only do the manual task, while (some of) the initial routine workers are (almost) indifferent between the routine and the abstract task. Then the wage gain of the manual workers will be $\triangle\pi_M$ and of the routine workers (almost) $\triangle\pi_A$. $\square$

B Details on the Derivation of Equations (7)–(12)

B.1 A Two Tasks Model for Illustration

First, another way of deriving Equation (7) without using the envelope theorem is illustrative: for simplicity consider a case with only two tasks, abstract and routine. Denote the specific value of the relative task price that makes worker i indifferent between the tasks as $\tilde{\pi}^i \equiv -(s_{Ai} - s_{Ri})$. For brevity, the indicator for choosing the abstract task is $I_{Ait} \equiv I_A(s_i, \pi_t) = 1[\pi_{At} + s_{Ai} > \pi_{Rt} + s_{Ri}] = 1[\pi_{At} - \pi_{Rt} > \tilde{\pi}^i]$ and the observed wage is $w_{it} = w_{Rit} + I_{Ait}(w_{Ait} - w_{Rit})$.

The change in i's wage $w_{i1} - w_{i0}$ when task prices change between $t = 0$ and $t = 1$ is therefore (remember $\triangle s_{Ai} = \triangle s_{Ri} = 0$):

$$\begin{aligned}
\triangle w_i &= \triangle\pi_R + I_{Ai1}(w_{Ai1} - w_{Ri1}) - I_{Ai0}(w_{Ai0} - w_{Ri0}) \\
&= \triangle\pi_R + \begin{cases} (\pi_{A1} - \pi_{R1}) - (\pi_{A0} - \pi_{R0}) = \triangle\pi_A - \triangle\pi_R & \text{if } I_{Ai0} = 1, I_{Ai1} = 1 \\ \pi_{A1} + s_{Ai} - (\pi_{R1} + s_{Ri}) = \pi_{A1} - \pi_{R1} - \tilde{\pi}^i & \text{if } I_{Ai0} = 0, I_{Ai1} = 1 \\ \pi_{R0} + s_{Ri} - (\pi_{A0} + s_{Ai}) = \tilde{\pi}^i - (\pi_{A0} - \pi_{R0}) & \text{if } I_{Ai0} = 1, I_{Ai1} = 0 \\ 0 & \text{if } I_{Ai0} = 0, I_{Ai1} = 0 \end{cases} && (16) \\
&= \triangle\pi_R + \int_{\pi_{A0}-\pi_{R0}}^{\pi_{A1}-\pi_{R1}} I_{Ait} d(\pi_{At} - \pi_{Rt}), && (17)
\end{aligned}$$

which is the two-task analog of (7). The step from (16) to (17) is an educated guess. Since the wage gain in (16) depends on the relative price range for which worker i chooses the abstract ($I_{Ait} = 1$) and routine ($I_{Ait} = 0$) tasks, respectively, this suggests an integral over the indicator function.

Check whether (17) is correct: if the worker always chooses abstract, he gets $\triangle\pi_A$ (first element of equation 16). If he never chooses abstract, he gets $\triangle\pi_R$ (fourth element of Equation 16). If he switches from routine to abstract, he gets

$$\triangle\pi_R + \int_{\pi_{A0}-\pi_{R0}}^{\tilde{\pi}^i} 0 d(\pi_{At} - \pi_{Rt}) + \int_{\tilde{\pi}^i}^{\pi_{A1}-\pi_{R1}} 1 d(\pi_{At} - \pi_{Rt}) = \triangle\pi_R + \pi_{A1} - \pi_{R1} - \tilde{\pi}^i,$$

which is the second element of Equation (16). Equivalently, if he switches from abstract to routine, he gets the third element of Equation (16).

Hence, since within tasks the wage gain is constant, the overall gain for a specific worker depends solely on the "distance" of the adjustment that the worker is still in the routine ($\tilde{\pi}^i - (\pi_{A0} - \pi_{R0})$) and already in the abstract ($\pi_{A1} - \pi_{R1} - \tilde{\pi}^i$) task. Notice that I never use the envelope theorem in this derivation. Also, the population distribution of skills does not appear here, as Equation (17) describes only one single worker's wage change when task prices change.

Taking expectations conditional on x_i in Equation (17) gives the two task version of result (9), suppressing the i_t sub-index used in the maintext for brevity:

$$E(w_{i1}|x_i) - E(w_{i0}|x_i) = \triangle\pi_R + \int_{\pi_{A0}-\pi_{R0}}^{\pi_{A1}-\pi_{R1}} p_A(x_i, \pi_{At} - \pi_{Rt}) d(\pi_{At} - \pi_{Rt}), \tag{18}$$

where $p_A(x_i, \pi_{At} - \pi_{Rt}) \equiv E[I_{Ait}|x_i] = Pr[\pi_{At} + s_{Ai} > \pi_{Rt} + s_{Ri}|x_i]$. Remember that s_{Ai}, s_{Ri} are a function of x_i (Equation 8 is an example). Now consider the linear interpolation (for brevity, $\tilde{\pi}_{ARt} \equiv \pi_{At} - \pi_{Rt}$)

$$p_A(x_i, \tilde{\pi}_{ARt}) \approx p_A(x_i, \tilde{\pi}_{AR0}) + [p_A(x_i, \tilde{\pi}_{AR1}) - p_A(x_i, \tilde{\pi}_{AR0})] \frac{\tilde{\pi}_{ARt} - \tilde{\pi}_{AR0}}{\tilde{\pi}_{AR1} - \tilde{\pi}_{AR0}}.$$

Plugging this interpolation into (18) and solving the integral yields the two task version of Equation (12):

$$E(w_{i1}|x_i) - E(w_{i0}|x_i) = \triangle\pi_R + \frac{p_A(x_i, \tilde{\pi}_{AR1}) + p_A(x_i, \tilde{\pi}_{AR0})}{2} \triangle(\pi_A - \pi_R)$$

B.2 The Three Tasks Model of the Main Text

This section now in detail derives the three task results from the main text using the envelope theorem. In fact Equation (7) is a little more subtle for the three-task case than the notation used in the main text reveals. For brevity, stick with the relative task price and skill notation employed at the end of the previous section so that $\tilde{\pi}_{KRt} \equiv \pi_{Kt} - \pi_{Rt}$ and $\tilde{s}_{KRi} \equiv s_{Ki} - s_{Ri}$ for $K\epsilon\{A, M\}$. Holding constant $\tilde{\pi}_{AR0}$ and $\tilde{\pi}_{MR0}$ and integrating Equation (6) from the main text with respect to π_{Rt} we get

$$w_i|_{\pi_{R1}, \tilde{\pi}_{AR0}, \tilde{\pi}_{MR0}} - w_i|_{\pi_{R0}, \tilde{\pi}_{AR0}, \tilde{\pi}_{MR0}} = \triangle\pi_R.$$

Similarly, integrating Equation (6) with respect to $\tilde{\pi}_{ARt}$ and $\tilde{\pi}_{MRt}$ but holding constant $\{\pi_{R1}, \tilde{\pi}_{MR0}\}$ and $\{\pi_{R1}, \tilde{\pi}_{AR1}\}$, respectively:

$$w_i|_{\pi_{R1},\tilde{\pi}_{AR1},\tilde{\pi}_{MR0}} - w_i|_{\pi_{R1},\tilde{\pi}_{AR0},\tilde{\pi}_{MR0}} = \int_{\tilde{\pi}_{AR0}}^{\tilde{\pi}_{AR1}} I_A(\tilde{s}_{ARi}, \tilde{s}_{MRi}, \tilde{\pi}_{ARt}, \tilde{\pi}_{MR0}) d\tilde{\pi}_{ARt}$$
$$w_i|_{\pi_{R1},\tilde{\pi}_{AR1},\tilde{\pi}_{MR1}} - w_i|_{\pi_{R1},\tilde{\pi}_{AR1},\tilde{\pi}_{MR0}} = \int_{\tilde{\pi}_{MR0}}^{\tilde{\pi}_{MR1}} I_M(\tilde{s}_{ARi}, \tilde{s}_{MRi}, \tilde{\pi}_{AR1}, \tilde{\pi}_{MRt}) d\tilde{\pi}_{MRt},$$

where the notation $I_A(\tilde{s}_{ARi}, \tilde{s}_{MRi}, \tilde{\pi}_{ARt}, \tilde{\pi}_{MR0})$ now makes explicit the dependence of $I_A(s_i, \pi_t)$ on relative skills and task prices (see also original Equation 3). Summing these three expressions gives Equation (7) from the main text with more detailed notation:

$$\triangle w_i = \triangle \pi_R + \int_{\tilde{\pi}_{AR0}}^{\tilde{\pi}_{AR1}} I_A(\tilde{s}_{ARi}, \tilde{s}_{MRi}, \tilde{\pi}_{ARt}, \tilde{\pi}_{MR0}) d\tilde{\pi}_{ARt} + \int_{\tilde{\pi}_{MR0}}^{\tilde{\pi}_{MR1}} I_M(\tilde{s}_{ARi}, \tilde{s}_{MRi}, \tilde{\pi}_{AR1}, \tilde{\pi}_{MRt}) d\tilde{\pi}_{MRt}.$$

Taking expectations conditional on x_i gives Equation (9), again suppressing the i_t subindex from the maintext:

$$\begin{aligned} E(w_{i1}|x_i) - E(w_{i0}|x_i) =& \triangle \pi_R + \int_{\tilde{\pi}_{AR0}}^{\tilde{\pi}_{AR1}} p_A(x_i, \tilde{\pi}_{ARt}, \tilde{\pi}_{MR0}) d\tilde{\pi}_{ARt} + \\ &+ \int_{\tilde{\pi}_{MR0}}^{\tilde{\pi}_{MR1}} p_M(x_i, \tilde{\pi}_{AR1}, \tilde{\pi}_{MRt}) d\tilde{\pi}_{MRt}, \end{aligned} \tag{19}$$

where

$$p_A(x_i, \tilde{\pi}_{ARt}, \tilde{\pi}_{MR0}) = Pr[\tilde{s}_{ARi} > -\tilde{\pi}_{ARt},\ \tilde{s}_{ARi} - \tilde{s}_{MRi} > -(\tilde{\pi}_{ARt} - \tilde{\pi}_{MR0})|x_i],$$

and similarly for $p_M(x_i, \tilde{\pi}_{ARt}, \tilde{\pi}_{MRt})$. Remember that (relative) skills are a function of x_i (Equation 8 is an example).

Linearly interpolating (19) by (dependence on x_i omitted for brevity)

$$p_A(\tilde{\pi}_{ARt}, \tilde{\pi}_{MR0}) \approx p_A(\tilde{\pi}_{AR0}, \tilde{\pi}_{MR0}) + [p_A(\tilde{\pi}_{AR1}, \tilde{\pi}_{MR1}) - p_A(\tilde{\pi}_{AR0}, \tilde{\pi}_{MR0})] \frac{\tilde{\pi}_{ARt} - \tilde{\pi}_{AR0}}{\tilde{\pi}_{AR1} - \tilde{\pi}_{AR0}}$$
$$p_M(\tilde{\pi}_{AR1}, \tilde{\pi}_{MRt}) \approx p_M(\tilde{\pi}_{AR0}, \tilde{\pi}_{MR0}) + [p_M(\tilde{\pi}_{AR1}, \tilde{\pi}_{MR1}) - p_M(\tilde{\pi}_{AR0}, \tilde{\pi}_{MR0})] \frac{\tilde{\pi}_{MRt} - \tilde{\pi}_{MR0}}{\tilde{\pi}_{MR1} - \tilde{\pi}_{MR0}}$$

gives Equation (12) from the main text. This contains another approximation on top of the interpolation insofar that one might prefer using $p_A(x_i, \tilde{\pi}_{AR1}, \tilde{\pi}_{MR0})$ instead of $p_A(x_i, \tilde{\pi}_{AR1}, \tilde{\pi}_{MR1})$ in the first approximation and $p_M(x_i, \tilde{\pi}_{AR1}, \tilde{\pi}_{MR0})$ instead of $p_M(x_i, \tilde{\pi}_{AR0}, \tilde{\pi}_{MR0})$ in the second, which are not observable in the data. The Monte Carlo simulations in Section 3.3 indicate that this is not a problem.

C Monte Carlo Simulations with Potential Confounders

The second part of the Monte Carlo simulations examines to what extent the Propensity Method is compromised when potentially confounding forces impact the wage distribution. A rigorous formal analysis of these confounders' impact on the estimator is outside the scope of this paper, so the Monte Carlos will be used to notify about potential threats to identification and adjustments to the Propensity Method that may account for them. I focus on three prominent such confounding forces here, rising returns to college, increasing college attainment, and a changing minimum wage. Table 6 reports the results of the Propensity Method with different adjustments, in order to save space only for the large sample (to precisely assess any bias of the estimator) and multivariate normality of the unobserved skills. The rest is the same as in Table 2.

The first two panels of Table 6 focus on changing returns to education, in particular to college attainment. Including college as a skill modifies potential wages (1) to

$$w_{Kit} = \pi_{Kt} + s_{Ki} + \lambda_t c_i \text{ with } K \epsilon \{A, R, M\}. \tag{1'}$$

c_i is a college dummy for individual i and λ_t the time-varying return to college. Assume first that c_i may be a function of observable or unobservable talents, but that the selection of talents into c_i does not change over time (i.e., it is not a function of task prices π_{Kt}). Then, Equation (12) becomes:

$$E(w_{i1} - w_{i0}|x_i) = \triangle\pi_R + \overline{p}_A(x_i)\triangle(\pi_A - \pi_R) + \overline{p}_M(x_i)\triangle(\pi_M - \pi_R) + E(c_i|x_i)\triangle\lambda \tag{12'}$$

One way to estimate these parameters is an augmented propensities regression (13):

$$\begin{aligned} w_{it} = {} & \alpha_0 + \alpha_1\overline{p}_A(x_i) + \alpha_2\overline{p}_M(x_i) + \alpha_3 \times 1[t=1] + \alpha_4\overline{p}_A(x_i) \times 1[t=1] + \\ & + \alpha_5\overline{p}_M(x_i) \times 1[t=1] + \alpha_6 c_i + \alpha_7 c_i \times 1[t=1] + \varepsilon_{it}, \end{aligned} \tag{13'}$$

where the task prices are identified by $\hat{\alpha}_3 = \widehat{\triangle\pi_R}$, $\hat{\alpha}_4 = \triangle(\widehat{\pi_A - \pi_R})$, $\hat{\alpha}_5 = \triangle(\widehat{\pi_M - \pi_R})$ again, and the change in the college premium by $\hat{\alpha}_7 = \triangle\lambda$. The first panel of Table 6 reports the results of this estimation when college is assumed to be dependent on observables x_i and some other random factors (e.g., credit constraints), but this relationship is not changing between period 0 and period 1.[34] The Table shows that not controlling for

[34]Specifically, I generate an index $c^* = f(x) + v_c = 9.0x_{math} + 4.8x_{verb} - 3.5x_{mech} + v_c$ with $v_c \sim N(0, 1 * s.d.[f(x)])$. The college indicator is set as $c = 1[c^* > Q_{c^*}(.75)]$, where $Q_{c^*}(.75)$ is the 75th per-

college overestimates the rising abstract task prices, because some of the rising college premium is attributed to that task. Including college in the wage regression as in (13′) solves this problem and also yields the correct college premium.

Table 6: Monte Carlo Simulations with College and Minimum Wage as Confounders

Confounder	**Adjustment**	$\triangle\pi_R$ **(s.e.m.)**	$\triangle(\pi_A - \pi_R)$ **(s.e.m.)**	$\triangle(\pi_M - \pi_R)$ **(s.e.m.)**	$\triangle\lambda$ **(s.e.m.)**
TRUE		-15.0	40.0	50.0	
Collg return rises	No adjustment	-12.36 (.48)	50.32 (.95)	48.39 (.72)	
$\triangle\lambda = 20.0$	2nd-stage collg	-14.50 (.48)	40.60 (.97)	49.38 (.71)	21.36 (.85)
Collg return occ-spec	2nd-stage collg	-14.59 (.48)	40.56 (.97)	49.40 (.72)	22.46 (.86)
& $\triangle\lambda = 20.0$	2nd-st collg X occ	-13.47 (.48)	41.37 (.94)	49.62 (.70)	19.18 (.64)
Collg attnm rises	No adjustment	-14.30 (.48)	41.16 (.94)	49.30 (.72)	
but $\lambda_1 = \lambda_0 = 0$	2nd-stage collg	-35.26 (.50)	72.58 (.94)	51.21 (.73)	-87.04 (.80)
Collg attnm rises	No adjustment	-2.98 (.48)	53.05 (.94)	46.52 (.72)	
& $\triangle\lambda = 20.0$	2nd-st pred collg	-1.27 (.01)	40.33 (.13)	49.35 (.95)	-7.74 (.03)
Min. wage rises	No adjustment	-13.71 (.49)	39.23 (1.06)	75.97 (.70)	
& no disempl.	Latent wage distr	-15.22 (.42)	39.73 (.94)	51.31 (.62)	

Notes: The table reports the mean estimated task prices (100 iterations) under different confounding factors and adjustments made to the Propensity Method in the sample with 40,000 individuals and normally distributed unobservables. The standard error of the mean estimate is shown in parentheses. For detail about parametrization and the assumed confounders, refer to the text and footnotes.

It is realistic to assume that returns to college differ across tasks. The second panel of Table 6 imposes that $\lambda_{Kt} = \lambda_K + \lambda_t$, with λ_K highest in the abstract and lowest in the routine task. This implies that workers gain from switching into tasks with higher college returns over time.[35] The second panel in Table 6 runs regression (13′) and an augmented version, allowing for different college returns in levels $\alpha_{K6}c_i$, $K\epsilon\{A, R, M\}$. For the purpose of identifying the task prices, regression (13′) performs well. The changing returns to college are unsurprisingly overstated because of the gains from switching. Regression (13′) also does not do worse considering the task prices than the version

centile of the c^* distribution. This matches well the cross-sectional relationship between c_i and x_i in the actual NLSY data. The college premium rises from $\lambda_0 = .2$ to $\lambda_1 = .4$.

[35]In particular, $\lambda_A = .4$, $\lambda_R = 0$, $\lambda_M = .2$, and $\triangle\lambda = .2$ as above. Complete flexibility of λ_{Kt} is not allowed because of the maintained assumption that the production of skills in tasks s_{Ki} is time-invariant and only the task prices change (see Section 2.3). The gain for a (college) worker from switching is $[(I_{Ai1} - I_{Ai0})(\lambda_A - \lambda_R) + (I_{Mi1} - I_{Mi0})(\lambda_M - \lambda_R)]c_i$.

accounting for $\alpha_{K6}c_i$. This indicates once more that correct modeling of wage *levels* in (13) or (13′) is not important for estimating the changing task prices, and that including college into the wage regression will suffice even when returns to college are task-specific.

The bigger challenge to the Propensity Method arises in the case of changing selection into college over time. From now assume that c_i is still dependent on x_i but that more individuals go to college in period 0 than in period 1.[36] The model in the third panel of Table 6 generates the (indeed plausible) case that the direct return to college is zero ($\lambda_t = 0,\ t\epsilon\{0,1\}$), and that all of the positive and rising college wage premium found in the data comes from it being associated with abstract tasks (i.e., the strict task model where worker characteristics are not priced directly). Intuitively, not controlling for college as in regression (13) should identify the correct task prices, because the true model return to college is zero. This is confirmed in the table, but when controlling for college the coefficient on $\triangle(\pi_{At} - \pi_{Rt})$ is severely upward-biased.[37] Therefore, the college dummy does not belong into the regression in this setting.

Of course, not including the college college dummy is also not an option when both, selection into college and the returns to college λ_t, change over time (fourth panel of Table 6). Neither controlling for college nor not controlling for college (not reported) work here. It turns out that empirically modeling the relationship between c_i and x_i and, just as with the task propensities, using predicted college $\hat{c}_i(x_i)$ identifies the relative task prices correctly. The reason is that the relationship between $\hat{c}_i(x_i)$ and x_i does not change over time other than a level shift in $\hat{c}_i(x_i)$ from period 0 to period 1. However, this depends on the assumed relationship between c_i and x_i here. In fact, additional unreported simulations show that if the college dummy c_i depends on unobserved skills u_i instead of x_i, this identification breaks down, while regressions (13) and (13′) work. My takeaway from these different settings is that the Propensity Method is at least partly able to account for rising returns to college and attainment. What seems important is to run both specifications, including and not including college, in regression (13) and

[36] Referring back to footnote 34, now $c = 1[c^* > Q_{c^*}(.75)]$ in period 0 and $c = 1[c^* > Q_{c^*}(.50)]$ in period 1.

[37] To see the difference between estimation of (13) and (13′) note that $\hat{\alpha}_4 = \frac{cov(w_{i1}, \tilde{\overline{p}}_A^{(1)}(x_i))}{var(\tilde{\overline{p}}_A^{(1)}(x_i))} - \frac{cov(w_{i0}, \tilde{\overline{p}}_A^{(0)}(x_i))}{var(\tilde{\overline{p}}_A^{(0)}(x_i))}$ where $\tilde{\overline{p}}_A^{(t)}(x_i) = \overline{p}_A^{(t)}(x_i)[1 - \frac{cov(\overline{p}_A^{(t)}(x_i), c_i^{(t)})}{var(c_i^{(t)})}]$ is the residual of regressing $\overline{p}_A^{(t)}(x_i)$ onto the college dummy and the other regressors in (13′) (omitted for brevity). If, as is the case here, $E(c_i^{(0)}|x_i^{(0)}) \neq E(c_j^{(1)}|x_j^{(1)})$ with $x_i^{(0)} = x_j^{(1)}$, then the comparability assumption 2 with a college regressor $\tilde{\overline{p}}_A^{(0)}(x_i) \neq \tilde{\overline{p}}_A^{(1)}(x_j)$ is violated, although without a college regressor $\overline{p}_A^{(0)}(x_i) = \overline{p}_A^{(1)}(x_j)$ it is not. Intuitively, comparability is violated because talent selection into the residual task propensity $\tilde{\overline{p}}_A^{(t)}(x_i)$ changes between the two periods, which induces a selection bias into $\hat{\alpha}_4$.

(13′) to see whether differences arise. Explicit modeling of the (endogenous) selection of observable skills into college may also be able to deal with this threat. But it is beyond the scope of this paper, which focuses on endogenous selection and changing returns to tasks, not college.

The last type of confounder considered is an increase in the minimum wage. This is of empirical relevance because the real minimum wage was raised substantially in the beginning of the 1990s. I follow Lee (1999)'s work here. In period 0 wages are censored at the 5th percentile of the overall wage distribution and in period 1 they are censored at the 10th percentile (*Case 1, "Censoring": no spillovers, no disemployment* in Lee, 1999). Lee estimates the latent wage distribution before censoring and adjusts workers' wages below the 50th percentile to match that latent distribution. The simulations do the same, assuming that the estimation recovers the true latent distribution. Note that workers do not get back their individual latent wages, only the overall distribution is matched. The last panel of Table 6 reports that the estimates for the manual task price change are severely overstated without the adjustment. However, when the minimum wage is correctly adjusted for, all the task price estimates are reasonably close to the truth again. More intricate effects of the minimum wage, like cases *2 "Spillovers"* and *3 "Truncation"* in Lee (1999), are beyond the scope of this paper.

D Detailed NLSY Sample Construction

I use data from the National Longitudinal Survey of Youth (NLSY) cohort of 1979 and 1997. The strength of the NLSY is that it provides detailed information about individuals' background and test scores in addition to education and labor market outcomes.

Individuals' labor market outcomes are evaluated at age 27 with the NLSY79 birth cohorts of 1957–64 reaching that age in 1984–92 and the NLSY97 birth cohorts of 1980–82 reaching it in 2007–09. Table 7 summarizes how the sample restrictions, attrition, and labor market participation reduce the sample size from 6,403 to 3,054 and from 4,599 to 1,207 males in the NLSY79 and the NLSY97, respectively. I restrict the sample to individuals who participated in the Armed Services Vocational Aptitude Battery of tests (ASVAB) in the first survey year. This restriction is necessary because the ASVAB provides measures of different dimensions of talent for each individual that are comparable over the two cohorts.

The participation in ASVAB is substantially lower in the NLSY97 than the NLSY79

Table 7: From the Full NLSY to the Analysis Sample

	NLSY79 (Birthyears 1956-1964)	NLSY97 (Birthyears 1980-1984)
Reason for exclusion		
Total males	6,403	4,599
Excluded oversampled white and older arrivers in U.S. than age 16	4,585	4,599
Birthyear > 1982	4,585	2,754
Type of attrition		
Ought to be present with ASVAB at age 27	4,585	2,754
No ASVAB excluded	4,299	2,081
%	94	76
Not present at age 27 excluded	3,939	1,737
%	86	63
Conditioned on working		
Excluded who report no or farm occupation, self-employed, and those with no wage income	3,054	1,207

Notes: The table reports how the analysis sample is constructed from the full NLSY 1979 and 1997, and where observations are lost or need to be dropped.

where almost everyone participated.[38] Moreover, sample attrition at age 27 is higher in the NLSY97 than the NLSY79 and overall only 63 percent of the NLSY79 participated in ASVAB and are also present at age 27. This problem is known (e.g. Altonji, Bharadwaj, and Lange, 2012; Aughinbaugh and Gardecki, 2007) and the attrition and non-test-participation rates in the data closely line up with those reported in the study by Altonji, Bharadwaj, and Lange (2012, henceforth ABL). The only difference is that ABL consider outcomes at the younger age of 22 and thus have slightly lower attrition rates.

In their paper, ABL note that the higher attrition rate in the NLSY97 may be partly due to NLSY97 respondents being first interviewed at ages 12–16 versus ages 14–21 for the NLSY79 and thus had more time to attrit. ABL further extensively examine the potential non-randomness of attrition and non-test-participation and its likely impact in

[38] According to the NLSY97 technical sampling report (Moore, Pedlow, Krishnamurty, and Wolter, 2000), nonrespondents to the ASVAB include ineligibles, refusals, breakoffs, and computer crashes, as well as individuals who are too ill or handicapped or with a language barrier. Moore, Pedlow, Krishnamurty, and Wolter (2000) find that this is higher among metropolitan youths, non-whites, males, and 16 year olds. They argue that there is a substantial impact of nonresponse only if the proportion of nonrespondents is high and if the differences between respondents and nonrespondents are high. Sampling weights, as used in this study, can account for differences in response rates between observable characteristics like the ones mentioned above.

biasing important labor market outcomes. Aughinbaugh and Gardecki (2007) do a similar exercise but focus on social and educational outcomes. Both studies find evidence that attrition is not random with respect to youths' outcomes and their backgrounds. However, Aughinbaugh and Gardecki (2007) conclude that attrition from the NLSY97 does not appear to affect inference when estimating the three outcomes at age 20 that they are considering and ABL decide that the differences between non-attriters and the whole sample are not forbidding.

Moreover, ABL carefully select the samples of NLSY79 and NLSY97 to make them comparable to one another and compute weights that adjust for attrition and non-test-participation on observable characteristics. I closely follow their procedures for constructing my own sample.[39] First, immigrants who arrived in the United States after age 16 are excluded from the NLSY79. This is done because the scope of the NLSY97 (age 12–16) also doesn't include older than age 16 arrivals. Second, I exclude the economically disadvantaged whites and military supplemental samples from the NLSY79 because they were discontinued early on in the survey and thus don't provide labor market outcomes at age 27 (or for ABL's purposes). Table 7 reports that 1,818 observations are dropped by making these restrictions to the sample. For each individual the observation that is closest to 27 years and 6 months of age is retained and labor market and final educational outcomes are measured from this observation.

ABL use a probit model to adjust the NLSY79 and NLSY97 base year sample weights to account for attrition and non-test-participation according to several observable characteristics, such as parental education, parental presence at age 14, indicators by birth-year, urban and SMSA residence status, indicator variables for race and gender, and an interviewer coded variable describing the attitude of the respondent during the interview. I also employ a probit model to adjust weights for attrition and non-test-participation and use the same specification and variables as ABL apart from leaving out parental presence at age 14. Alternatively, a fully stratified set of indicators for birthyear, year, sex, and race, as employed by the Bureau of Labor Statistics for weighting, yields very similar results.[40] As ABL do in their paper, I proceed from this point with the assumption that, after attrition weighting, the two NLSY samples are representative of the population of young Americans that they are supposed to cover. These samples have the size of 3,939

[39]Thus, for more information on the sample construction and for statistics on the effects of attrition, please refer to ABL in addition to the description provided here. I would like to thank Prashant Bharadwaj for providing me with their data and do-files.

[40]I thank Steve McClaskie and Jay Zagorsky for providing me with the official attrition-adjusted sample weighting program for the NLSY.

and 1,737 individuals in the NLSY79 and the NLSY97, respectively.

I follow Lemieux (2006), who uses CPS Outgoing Rotation Group data, in how I compute wages and in defining the sample of working individuals (henceforth labor supply). Hourly wages reported for the current main job are used and normalized to 1979 real values by adjusting with the PCE deflator provided by the St.Louis Federal Reserve Bank.[41] While Lemieux (2006) removes outliers with 1979 real hourly wages below $1 and above $100, I remove the high wages from $40 onward because the NLSY wage data is very inaccurate for values above this threshold.

Finally, in order to condition on working individuals, all individuals who report not to be self-employed, and who are employed in a non-farm, non-fishing and non-forestry occupation according to the Census 1990 three-digit occupation classification are left in the sample. This leaves me with an analysis sample of 3,054 and 1,207 males in the NLSY79 and NLSY97, respectively (compare table (7) again). As in Lemieux (2006), all of those individuals are weighted by the number of hours that they work per week on top of the sample weights that are adjusted for test-participation and attrition.

E An Alternative Dataset: the Census/ACS Sample

This section repeats the estimation of changes in task prices and the assessment of their effect on the overall wage distribution in the Census/ACS sample. In addition, Acemoglu and Autor (2011) have used the Census/ACS to examine the wage growth of starters in abstract and manual tasks versus routine tasks similar to Section 5.1.

E.1 Data Construction and Identification Assumptions

The Census/ACS is from Section 5 of Acemoglu and Autor (2011)'s chapter in the Handbook of Labor Economics. Acemoglu and Autor construct a dataset of 16 to 64 year old workers from the decennial censuses and the American Community Survey, which they split into demographic cells indexed by education, age, and region.[42] They then examine the wage changes of the demographic cells in relation to their initial specializations in abstract, routine, and manual tasks in 1959 separately by gender and by decade. The tasks are again the broad occupation categories described in Section 4.1 of this paper. I focus on

[41] Source: "Personal Consumption Expenditures: Chain-type Price Index (PCECTPI)", accessed 2012-8-14, http://research.stlouisfed.org/fred2/series/PCECTPI

[42] Data are downloaded from David Autor's website http://economics.mit.edu/faculty/dautor/data/acemoglu, accessed 2014-12-15. Individuals need to report having worked last year but military and agricultural workers are excluded.

Acemoglu and Autor's data for males here, while the results for females are summarized in Appendix F. The labor market facts in the Census/ACS data are qualitatively similar to the ones plotted in Figure 2 for the NLSY.[43]

The data from Acemoglu and Autor (2011) is an attractive complement to the NLSY sample for the task price estimation because it alleviates a couple of concerns that one might have with the NLSY. In the NLSY sample the analysis is conducted for the periods 1984–1992 versus 2007–2009. This may confound the effect of RBTC with the business cycle, as the years 2008 and 2009 are part of the great recession. In contrast, in the Census/ACS data, the task price changes can be estimated for the comparable period 1989–2007 as well as the separate sub-periods 1979–1989, 1989–1999, 1999–2007. Further, the NLSY data is based on 27 year olds, while in the Census/ACS I can examine a large and representative sample of all workers age 16–64 and compare the results.

The first-stage variation in the Census/ACS data are average abstract and manual task propensities for 80 demographic cells, which, due to the large sample size, can be computed non-parametrically as simple cell averages. Table 8 reports for each sample year the shares among male employment of the underlying interacted five education, four region, and four age groups. The table shows that male workers' educational attainment rises, that they are aging, and that they are moving to the South and West over time. However, the regional and especially the educational trends appear quite modest after 1989, which is consistent with the findings in Autor (2015), among others, that young males' educational attainment has increased remarkably little since the 1980s. The changing overall age structure is predetermined by birth rates decades earlier and hardly a reaction to RBTC. Therefore, these moderate trends may not necessarily indicate a critical violation of the comparability Assumption 2, especially for the estimation periods after 1989.

In addition, inspection of Equation (13) reveals that adding a constant to the regressors does not change the regression coeffients α_4 and α_5 (α_3 is affected, though). This allows for the following ad-hoc generalization of the comparability assumption:

Assumption 3 (Comparability-in-differences)**.** *Differences between individuals i and l with*

[43] Figure 13c in Acemoglu and Autor (2011)'s paper shows that there was job polarization among males in the Census/ACS data during 1989–2007. I reproduce this finding in my own computations (unreported). I further find that wages in the manual and the routine task stagnate during 1989–2007, while wages in the abstract task rise strongly. Therefore, there is again no clear evidence of wage polarization in tasks. Figure 9 in Acemoglu and Autor (2011) shows that hourly wages polarized in an additional sample that they construct from the CPS ORG during 1988–2008. Acemoglu and Autor (2011)'s Figure 9b is reproduced in Panel (a) of Figure 5 below, and it is qualitatively comparable to the bottom panel of Figure 2 for 27 year olds in the NLSY data.

Table 8: Male Employment in the Census/ACS with Respect to Education, Age, and Region by Year

	1979	**1989**	**1999**	**2007**
Nbr of observations ('000)	2,589	2,834	3,158	667
Share of Education Group (%)				
High School Dropout	25	15	13	12
High School Graduate	32	33	31	30
Some College	23	29	30	31
4 year College	12	15	17	18
Graduate Degree	7	8	9	9
Share of Age Group (%)				
16-24	27	22	19	20
25-39	38	42	38	35
40-54	24	26	33	33
55-64	11	10	9	12
Share of Region (%)				
Northeast	22	21	19	18
Midwest	27	24	24	23
South	32	33	35	35
West	19	21	22	23

Notes: The table shows education, age, and region groups' shares of employment in the Census/ACS sample as constructed by Acemoglu and Autor (2011). A full interaction of these groups yields 80 demographic cells for the task price estimation.

the same vectors x_{i_t} and x_{l_t} are comparable over time. That is, for all x_{i_0}=x_{i_1}, x_{l_0}=x_{l_1} ($i_0 \neq i_1 \neq l_0 \neq l_1$), the population distributions of relative unobservable skills between i and l, $u_{Ki_0} - u_{Kl_0} = \delta_K'(z_{i_0} - z_{l_0})$ and $u_{Ki_1} - u_{Kl_1} = \delta_K'(z_{i_1} - z_{l_1})$, $K\epsilon\{A, R, M\}$, are the same.

In addition, the relative task choice probabilities are also approximately constant over time.[44] *That is, for all x_{i_0}=x_{i_1}, x_{l_0}=x_{l_1}, and $\pi_t\epsilon\{\pi_0, \pi_1\}$,*

$$p_A(x_{i_0}, \pi_t) - p_A(x_{l_0}, \pi_t) \approx p_A(x_{i_1}, \pi_t) - p_A(x_{l_1}, \pi_t)$$
$$p_M(x_{i_0}, \pi_t) - p_M(x_{l_0}, \pi_t) \approx p_M(x_{i_1}, \pi_t) - p_M(x_{l_1}, \pi_t).$$

As in Sections 3.1–3.2, a sub-index i_t indicates an individual i observed in the data in t. Under Assumption 3, the correct relative task prices $\triangle(\pi_A - \pi_R)$ and $\triangle(\pi_M - \pi_R)$, but not the level $\triangle\pi_R$, can be identified from regression (13) even if the unobservable

[44]They need not be exactly constant, as the probabilities (10) are nonlinear functions of their arguments.

skill selection into observable groups x_{i_t} is changing over time.

To see this better, suppose workers with changing unobservable talent z_{i_t} enter different education groups x_{i_t} over time. The relative task price estimates from such data may still be correct as long as the shifts in unobservable talent and task propensities are similar across x_{i_t} groups. For example, this is the case if all x_{i_t} groups become equally less skilled in abstract tasks over time (e.g., because the most talented of a given education group become the least talented of the next higher education group), and $p_A(x_{i_1}, \pi_t) \approx p_A(x_{i_0}, \pi_t) + c_{A01}$ for all $x_{i_0} = x_{i_1}$ and $c_{A01} < 0$ (similarly for the manual task).

Consistent with Assumptions 2 and 3, an instrumental variables strategy carried out below, which removes potential unobservable talent selection effects from task propensities, yields similar task price estimates as the baseline OLS specification. In another robustness check, I directly control for education, age, and region main effects to exploit only the differential variation across demographic cells with the same levels of these variables. Again, the estimated task prices for the joint 1990s and 2000s, which are comparable to the NSLY and long enough for RBTC—if important—to dominate alternative forces, hardly change.

To conclude, while the identification assumptions are somewhat more debatable in the Census/ACS compared to the NLSY, it should still increase confidence in the results if both datasets arrive at similar estimates.

E.2 Task Price Estimates and Effect on the Overall Wage Distribution

This section estimates the task price changes for all males in the Census/ACS over the joint 1990s and 2000s as well as by decade. Their effect on the overall wage distribution is also assessed.

Panel A in Table 9 reports task price estimates under the same baseline specification as in row one of Table 5. During the 1980s (1979–1989), the prices for the routine as well as the manual tasks declined substantially while the price for the abstract task increased. In contrast, during the 1990s, manual task prices increased strongly and overall task prices polarized. The trend then reversed in the 2000s before the great recession (1999–2007), when only the price for the abstract task rose. This is in line with the findings in the literature that, despite strong employment gains in the manual task, the bottom of the wage distribution did not rise during this period (Acemoglu and Autor, 2011; Mishel, Shierholz, and Schmitt, 2013).

Table 9: Estimated Task Price Changes in the Census/ACS (Different Periods)

		$\triangle(\pi_A - \pi_R)$ in log points (s.e.)	$\triangle(\pi_M - \pi_R)$ in log points (s.e.)	$\triangle\pi_R$ in log points (s.e.)
Panel A:	1979-1989	30.5 (4.6)	-1.2 (17.9)	-16.8 (3.2)
OLS on Demogr. Cells	1989-1999	17.6 (3.4)	54.2 (13.2)	-8.5 (3.3)
(Baseline)	1999-2007	14.1 (2.8)	-15.0 (10.0)	-8.7 (2.2)
	1989-2007	**33.6 (3.2)**	**41.3 (11.2)**	**-19.0 (2.4)**
Panel B:	1979-1989	36.1 (4.8)	21.9 (19.2)	-20.9 (3.4)
OLS on Demogr. Cells	1989-1999	16.5 (3.8)	51.6 (14.7)	-7.9 (2.7)
(IV using 1959&69 Propnsts)	1999-2007	15.2 (3.1)	-10.3 (11.4)	-9.7 (2.5)
	1989-2007	**33.6 (3.4)**	**43.2 (12.8)**	**-19.2 (2.7)**
Panel C:	1979-1989	3.5 (19.9)	25.2 (21.4)	-18.1 (5.7)
OLS on Demogr. Cells	1989-1999	23.8 (14.9)	23.8 (15.7)	-3.8 (4.7)
(Educ, Age, Region Cntrls)	1999-2007	12.9 (8.2)	32.0 (10.0)	-23.0 (3.3)
	1989-2007	**42.3 (16.0)**	**49.8 (19.0)**	**-25.4 (6.2)**
Panel D:	1979-1989	40.1 (4.5)	82.6 (17.5)	-26.9 (3.1)
OLS on Demogr. Cells	1989-1999	16.4 (3.4)	42.0 (12.7)	-7.1 (2.5)
(Adj. for min. wage)	1999-2007	14.7 (2.8)	-9.9 (10.0)	-9.5 (2.2)
	1989-2007	**33.0 (3.2)**	**34.3 (11.0)**	**-18.1 (2.4)**

Notes: Panel A of the table presents estimated task price changes for the baseline Propensity Regression (13) in different time periods. Panel B instruments the task propensities with their values from before the sample period (1959 and 1969). Panel C controls for education, age, and region main effects. Panel D reports baseline estimates when wages are first adjusted for the change in the real value of the minimum wage as in Lee (1999). Standard errors in parentheses next to the coefficients.

The last row of Panel A reports the results for the joint 1990s and early 2000s (1989–2007), which is broadly comparable to the period examined in the NLSY sample without the great recession part. During that period, task prices overall polarized substantially and by about one third more strongly than in the NLSY sample. That the effect is stronger for the Census/ACS seems plausible, since the on average more experienced workers in that sample are probably less able to adjust their tasks than the young workers in the NLSY. In addition, the effects on all task prices are now clearly statistically significant.

If the comparability-in-differences Assumption 3 were violated in Panel A of Table 9, the changing unobservable talent selection into a given education-age-region cell would lead to counterfactual task propensities that deviate systematically from the actual task propensities that they are supposed to capture. The instrumental variables estimation in Panel B extracts good variation from the task propensities, removing some potential bad variation that is due to changing selection of unobservable talent into education-age-region cells. Practically, I instrument for all $p_K(x_i, \pi_t)$ using occupational propensities in the periods before the estimation sample (1959 and 1969), which leads to regressors entering Propensity Regression (13) that contain the time-invariant comparative advantage of x_i cells in the abstract and manual tasks.[45] With the exception of the manual task in the 1980s, the task prices in Table 9, Panel B are close to those in Panel A. This suggests that differentially changing unobservables selection into education-age-region cells does not play a large role in the estimates of Panel A, and that comparability-in-differences may not be critically violated in the Census/ACS data.

Panel C of Table 9 reports a different robustness check for the task price estimates, controlling for education, age, and region main effects. The identifying variation that is left with these controls are the differential task propensities across demographic cells for given levels of education, age, and region. Therefore, this specification removes violations of the comparability-in-differences assumption that occur due to overall shifts in the demographic composition of the population. It also removes the effects of alternative forces to RBTC that are based on across-the-board changing demand for specific education, age, or region groups (e.g., SBTC).

A limitation of this specification is that, with the main effects included, there is not a lot of variation remaining for identification. This is reflected in the first three rows of Panel C, where the relative task prices are insignificant for all but one of the decadal changes (qualitatively, task prices still polarize during the 1990s). However, it is supportive of Panels A and B that the task price estimates for the long period of 1989–2007, where the impact of RBTC should dominate alternative forces in the data, are similar and remain statistically significant. Consistent with these arguments, a Hausman test does not reject the restricted estimation model (Panel A) against the unrestricted model (Panel C) in any of the four periods (details available upon request).

[45]For example, in the case of the abstract task under $t = 0$ prices, a violation of assumption 3 implies that $p_A(x_{i'}^{(1)}, \pi_0) = p_A(x_i^{(0)}, \pi_0) + c_{A01} + g(x_i^{(0)}, x_{i'}^{(1)})$, where $x_{i'}^{(1)} = x_i^{(0)}$ and $g(.)$ is a mean-zero non-constant function. Further, when the abstract task price increases from $t = 0$ to $t = 1$, workers with a given x_i should become more likely to enter that task. Formally, $p_A(x_i^{(t)}, \pi_1) = f(p_A(x_i^{(t)}, \pi_0))$, with $f'(.) \geq 0$. Assuming $f(.)$ to be approximately linear obtains the actual $t = 1$ task propensity as a weighted sum of the actual $t = 0$ task propensity and a systematic error due to changing unobservable talent selection into $x_i^{(t)}$ over time,

$$p_A(x_{i'}^{(1)}, \pi_1) \approx \phi_1 + \phi_2 p_A(x_i^{(0)}, \pi_0) + \phi_3 g(x_i^{(0)}, x_{i'}^{(1)}) \text{ with } \phi_2 \geq 0.$$

The instrumental variables estimation in Panel B of Table 9 extracts good variation $\hat{\phi}_1 + \hat{\phi}_2 p_A(x_i^{(0)}, \pi_0)$ from the propensity $p_A(x_i^{(1)}, \pi_1)$, removing some of the $\phi_3 g(x_i^{(0)}, x_{i'}^{(1)})$ variation that is due to changing selection of unobservable talent into education-age-region cells.

The last Panel of Table 9 reports the task price estimates when wages are adjusted for the increase in the real value of the minimum wage analogous to row five in Table 5. As in that latter table, the adjustment for the real value of the minimum wage does not have a large effect on the task price estimates for the 1990s and 2000s. The only difference is the changing and high increase of the manual task price during the 1980s. This may be related to Lee (1999)'s finding that the wage distribution (for males and females) would have polarized during the 1980s without the decline in the real value of the minimum wage.[46]

Overall, there exists quite robust evidence for task price polarization over the joint 1990s and 2000s in the Census/ACS, too. This result substantiates RBTC's importance during that period beyond the existing, and at first glance partly contradicting, evidence on the occupational and overall wage distributions (see discussion in Section 2.2). It also lends support to the estimation method for task prices, which derives its identification from the interplay between workers' sorting into tasks and their wage growth.

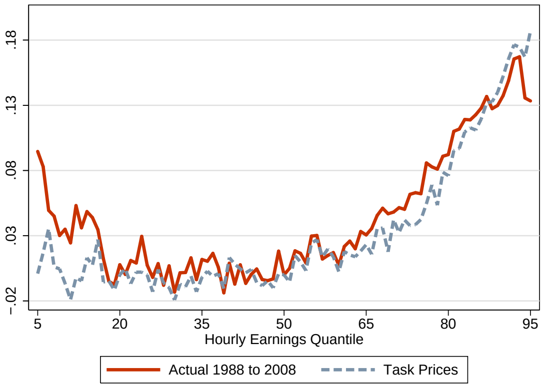
Finally, I assess the effect that changes of task prices in the Census/ACS may have had on aggregate wage inequality. As in Equation (15), a predicted wage distribution is generated by assigning every worker the price estimate for his initial task. In Panel (a) of Figure 5, I use the task price estimates from the baseline specification in Table 9 to plot the predicted wage distribution into Figure 9b of Acemoglu and Autor's Handbook chapter (Figure 9b uses CPS data for males of all ages and the period is 1988–2008 while the task price estimates are from the Census/ACS for 1989–2007). The resulting predicted wage distribution closely follows the actual at the top, in the middle, and almost to the lower end of the distribution. Only at the very bottom is the predicted wage distribution again lower than the actual as for the 27 year olds. Adjusting for the increase in the minimum wage analogous to Panels (c) and (d) of Figure 4, closes part of this gap in the bottom (Figure 5, Panel b).

In the Census/ACS, task prices' contribution to the evolution of the wage distribution can also be examined by decade. Using the estimates for the sub-periods in Panel A of

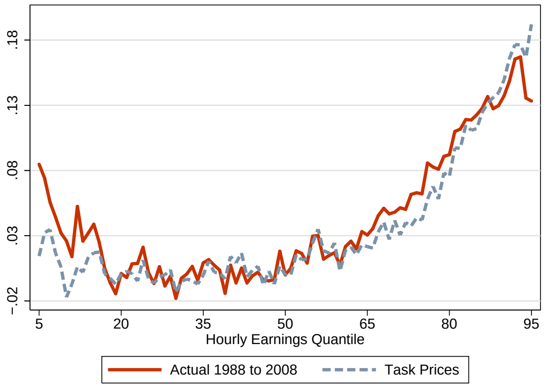
[46]More generally, the task price estimates for the 1980s are quite sensitive to the different specifications in Panels A–D of Table 9. This could be due to a couple of forces that may have compromised comparability-in-differences or worked aside from RBTC during that period. First, the substantial decline in the real value of the minimum wage seems to be responsible for the difference between Panels A and D. In addition, previous literature has found a strong influence of SBTC on wages and a rapidly rising college premium during the 1980s (e.g., Acemoglu and Autor, 2011). This may have affected the labor market separate from RBTC, leading to the difference between Panels A and C. Educational attainment also rose rather substantially during the 1980s (see Table 8). This may have changed the unobservable talent composition within education-age-region cells and affected the comparison between Panels A and B.

Figure 5: Actual and predicted change in the wage distribution due to Census/ACS task prices (with and without adjustment for minimum wage and by decade)

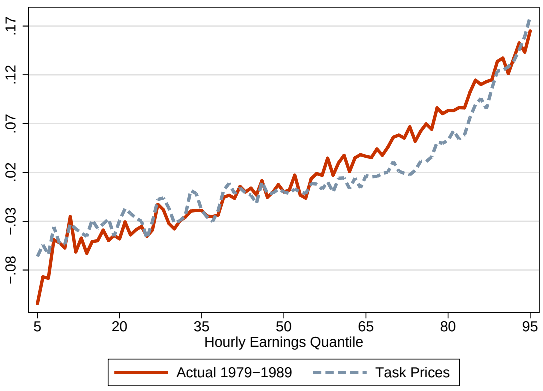
(a) CPS (as in Acemoglu and Autor, 2011, Figure 9b)



(b) CPS, minimum wage adjusted



(c) 1979–1989



(d) 1989–1999

(e) 1999–2007

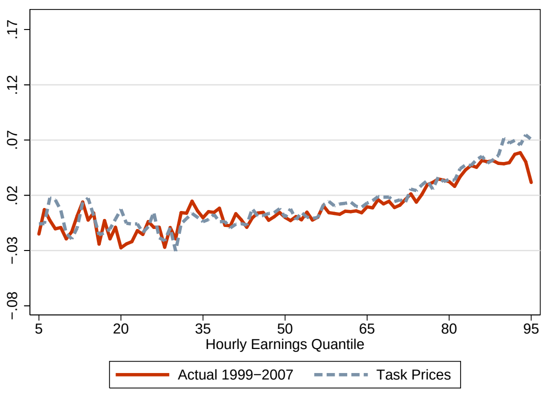


Table 9, I plot the actual and the predicted distributions for all males in 1979–1989, 1989–1999, and 1999–2007. During the 1980s, the fit is decent (Figure 5, Panel (c)). Inequality rises across-the-board in the predicted as well as the actual wage distribution, with the predicted dropping less in the very bottom and rising somewhat less between the 50th to 90th percentile. During the 1990s and the 2000s, the fit is good. In both the actual and the predicted distribution, wages are polarizing in the 1990s and the two lines essentially overlap everywhere except the very bottom. In 1999–2007, both, the actual and the predicted wage distribution, are largely flat up to the 60th percentile, after which they grow continuously together. Therefore, the task prices by themselves do well in explaining the polarization of the wage distribution by decade and overall during the 1990s and 2000s before the great recession.

F Task Price Estimates for Females

This section reports the task price estimates for females. As in the main text, regression 13 is run on predicted task probabilities from a first-stage multinomial logit in the NLSY sample and on actual choice frequencies for discrete demographic cells in the Census/ACS sample.

Table 10 reports the results. In the first row of Panel A the task prices for 27 year old females between the NLSY79 and the NLSY97 move in the opposite direction of what is predicted by RBTC. The relative price for the abstract and manual task fall substantially, while the price for the routine task rises. The same happens when college dummies are controlled for in the second stage wage regression (second row of Panel A).

The reason for this unexpected result should be that the identification Assumptions 1–3 are violated or that RBTC is dominated by other forces for females. One identification violation that was already mentioned in the main text is on the first-stage Assumption 1, as in the (unreported) multinomial logit regression the mechanical talent for females is unrelated to task choice in the NLSY79, but it predicts that the abstract task will not be chosen in the NLSY97. Therefore, the three talents do not strongly and consistently predict task choices in the two NLSYs for females.[47] Row three of Panel A supports this suspicion, as the direction of the task price estimates reverses and moves in the initially expected direction when a college dummy is included in the first-stage choice regression. Overall, the results on the task prices in the NLSY are therefore inconclusive.

[47] The pseudo R-squared in the choice regressions for females is also substantially lower than for males in

Table 10: Estimated Task Prices in the NLSY and Census/ACS Sample—Females

		$\triangle(\pi_A - \pi_R)$ in log points (s.e.)	$\triangle(\pi_M - \pi_R)$ in log points (s.e.)	$\triangle \pi_R$ in log points (s.e.)
Panel A (NLSY): 1984/92 to 2007/09	Baseline	-52.9 (21.4)	-88.7 (38.8)	48.7 (14.1)
	2nd-stage college	-60.6 (21.1)	-53.9 (36.5)	36.9 (13.2)
	1st-stage college	19.1 (15.1)	26.9 (32.8)	-2.1 (10.9)
Panel B (Census/ACS): OLS on Demogr. Cells (Baseline)	1979-1989	12.5 (5.2)	-28.6 (12.0)	2.2 (3.5)
	1989-1999	4.8 (3.5)	2.2 (8.1)	8.2 (2.6)
	1999-2007	1.5 (4.2)	-23.4 (8.9)	-0.8 (3.3)
	1989-2007	**6.2 (5.2)**	**-23.4 (11.4)**	**-7.4 (4.0)**
Panel C (Census/ACS): OLS on Demogr. Cells (Educ, Age, Region Cntrls)	1979-1989	14.8 (13.4)	12.1 (15.6)	-8.4 (6.0)
	1989-1999	-2.8 (12.4)	-25.3 (14.0)	15.5 (5.9)
	1999-2007	-23.5 (9.7)	-43.6 (9.3)	6.8 (4.6)
	1989-2007	**-17.3 (16.2)**	**-63.3 (16.1)**	**19.6 (7.5)**

Notes: Panel A of the table presents estimated task price changes for females in the NLSY sample. The two rows correspond to the specifications reported in the first two rows in Table 5 for males. The third row includes a college dummy into the first-stage choice regression. Panels B and C present task price change estimates in the Census/ACS sample for females baseline and with main effects for education, age, and region category, respectively. These correspond to Panels A and C of Table 9 for males. Standard errors in parentheses next to the coefficients.

Panels B and C of Table 10 report the results for females in the Census/ACS sample, which are again not clearly supportive of RBTC. First, the estimated relative price of the abstract task rises by around 12–15 log points during the 1980s, but it stagnates (Panel B) or even falls when controlling for main effects (Panel C) in the 1990s and 2000s. Moreover, the relative price for the manual task falls substantially in both specifications over the last two decades. Only in the 1980s, when controlling for main effects in Panel C, do the estimated relative prices for the abstract and the manual task rise as is predicted by RBTC.[48]

The results reported for females in Table 10 as well as for males in Table 9 are in line with Acemoglu and Autor (2011). Using the same data as in Table 10 Panels B and C, Acemoglu and Autor (2011) find that, while for males initial specialization in abstract and manual tasks predicts higher wage growth over the last decades, the results for females are rather inconclusive. As explained in Section 4.1, the author's view is that this is because the demand and supply of skills for females has changed quite drastically for reasons which come on top of RBTC or which even dominate it during this period.

Table 12.

[48]All of the task price estimates in panels A–C are similar when adjusting for the change in real value of the minimum wage as in the main text.

G Additional Tables and Figures

Table 11: Male Employment in the NLSY with Respect to Average Demographics, Early, and Contemporary Skill Determinants

	NLSY79	NLSY97
Nbr of observations	3,054	1,207
Percentage of observations	71.60	28.40
Demographics		
Age	27.00	27.00
White	0.80	0.72
Black	0.13	0.14
Hispanic	0.06	0.14
Early skill determinants		
AFQT	167.31	167.65
Low AFQT Tercile	0.34	0.33
Middle AFQT Tercile	0.33	0.34
High AFQT Tercile	0.33	0.32
Math Score (NCE)	50.45	50.73
Verbal Score (NCE)	50.26	50.49
Mechanical Score (NCE)	50.41	50.69
Illicit Activities (NCE, Measured 1980)	49.98	50.01
Precocious Sex (NCE, Measured 1983)	49.91	50.24
Mother's Education (Years)	11.86	13.11
Father's Education (Years)	10.83	13.09
Contemporary skill determinants		
High School Dropout (HSD)	0.12	0.07
High School Graduate (HSG)	0.43	0.58
Some College (SC)	0.20	0.06
College Graduate (CG)	0.19	0.24
Advanced Degree (AD)	0.06	0.04
North East	0.22	0.17
North Central	0.29	0.25
South	0.32	0.35
West	0.17	0.21

Notes: The table shows average demographics and skill proxies in the NLSY79 and NLSY97 for all males weighted by hours worked. NCE indicates variables in the population (including non-workers) are standardized to "normal curve equivalents" with mean 50 and standard deviation 21.06. This is done when absolute values of these variables cannot be compared over the two cohorts.

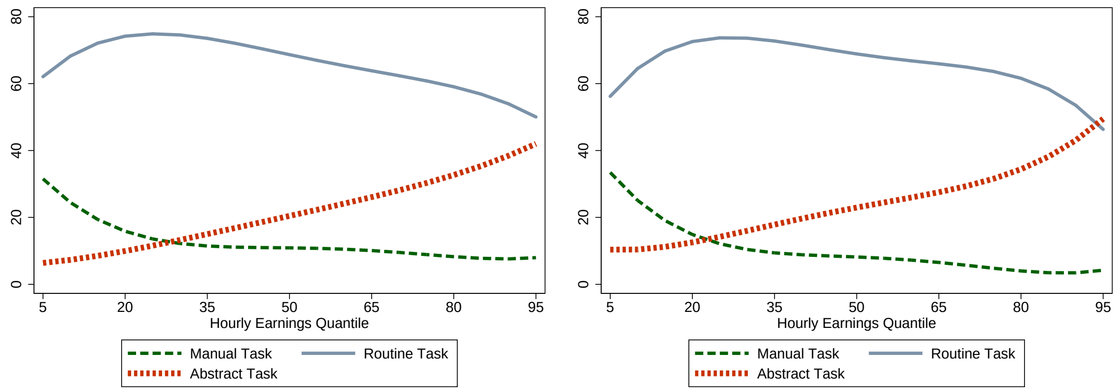
Table 12: Sorting into Tasks in the NLSY, Multinomial Logit Regressions

	(1) NLSY79	(2) NLSY79	(3) NLSY97	(4) NLSY97
Abstract				
Constant	-4.024*	-1.710*	-3.176*	-1.384*
Black	0.235	0.159	-0.152	-0.106
Hispanic	0.03	-0.031	-0.472	-0.456
Math (NCE)	0.047*		0.034*	
Verbal (NCE)	0.023*		0.032*	
Mechanic (NCE)	-0.014*		-0.019*	
Middle Math Tercile		1.144*		0.441
High Math Tercile		2.315*		1.426*
Middle Verbal Tercile		0.207		0.670*
High Verbal Tercile		0.750*		1.445*
Middle Mechanic Tercile		-0.269		-0.258
High Mechanic Tercile		-0.552*		-0.618*
Illicit Activities (NCE)		-0.009*		-0.003
Precocious Sex (NCE)		-0.004		-0.006
Manual				
Constant	-1.689*	-1.608*	-1.339*	-2.053*
Black	0.636*	0.762*	0.473*	0.658*
Hispanic	0.201	0.243	-0.216	-0.114
Math (NCE)	-0.002		-0.009	
Verbal (NCE)	0.018*		0.021*	
Mechanic (NCE)	-0.023*		-0.017*	
Middle Math Tercile		-0.381*		-0.07
High Math Tercile		0.128		-0.395
Middle Verbal Tercile		0.342		0.27
High Verbal Tercile		0.471		0.790*
Middle Mechanic Tercile		-0.319		-0.281
High Mechanic Tercile		-0.908*		-0.608
Illicit Activities (NCE)		-0.002		0.013
Precocious Sex (NCE)		-0.003		-0.003
Pseudo R-Squared	0.132	0.123	0.114	0.112
N	2936	2936	1210	1210

Notes: Each column presents the results from a multinomial logit regression of task choice on demographics and talent proxies. The omitted group is the routine task. The first column uses only linear test scores in the NLSY79. The second column, which is the specification to estimate task propensities in the following, uses terciles of test scores and adds measures of risky behavior. The last two columns repeat these estimations for the NLSY97. Not to overload the table, significance at the five percent level is indicated by a single *.

Figure 6: Representation of Tasks in the Wage Distribution, 1979 Cohort

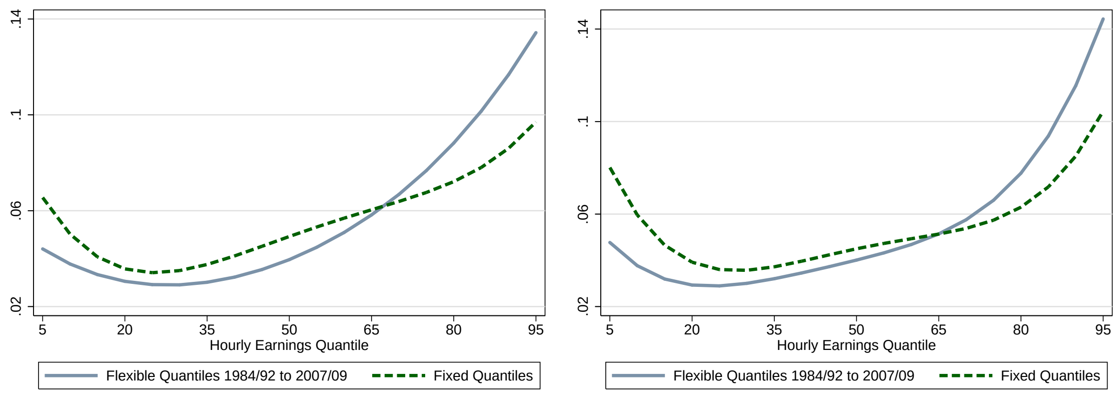
(a) NLSY, males age 27 (b) CPS, males age 27



Notes: The figure plots the smoothed employment share of the manual, routine, and abstract tasks within the quantiles of the wage distribution. Smoothing is done using the predicted values from a fourth order polynomial regression of the employment shares on the quantiles.

Figure 7: Smoothed predicted change in the wage distribution due to changing task prices, flexible and fixed quantiles

(a) NLSY, males age 27 (b) CPS, males age 27



Notes: The solid line depicts the predicted change in log real wages along the quantiles of the wage distribution due to estimated changes in task prices. The dashed line depicts the same predicted change when individuals are fixed at their original quantiles in the wage distribution. The lines are smoothed because for the predicted under fixed quantiles the individuals who correspond to these quantiles exclusively determine their change. This would make the predicted change very spiky. Smoothing is done using the predicted values from a fourth order polynomial regression of average wage changes on the quantiles.