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

Michael J. Böhm *

October 2019

Abstract: This paper proposes a new approach to estimate task prices per efficiency unit of skill in the Roy model. I show how the sorting of workers into tasks and their associated wage growth can be used to identify changes in task prices under relatively weak assumptions. The estimation exploits the fact that the returns to observable talents will change differentially over time depending on the changes in prices of those tasks that they predict workers to sort into. In the generalized Roy model, also the average non-pecuniary amenities in each task are identified. I apply this approach to the literature on routine-biased technical change, a key prediction of which is that task prices should polarize. Empirical results for male workers in U.S. data indicate that abstract and manual tasks' relative prices indeed increased during the 1990s and 2000s.

Keywords: Task Prices; Roy Model; Routine-Biased Technical Change; Polarization; Wage Distribution
JEL Classification Numbers: J23, J24, J31

* University of Bonn and Institute of Labor Economics (IZA). email: 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, Felix Schran, and Christian Siegel as well as participants at various seminars for very helpful comments. I would also like to thank the anonymous referees and especially the editor, Chris Taber, for requesting revisions that much improved the paper.

1 Introduction

The prices that are paid per efficiency unit of skill in different occupations or industries are a key quantity in labor economics. For example, a classic paper by Heckman and Sedlacek (1985) estimates the changes of such task prices in the manufacturing and non-manufacturing sectors of the United States during 1968–81 (see also Young, 2014). In the more recent literature on routine-biased technical change, Cortes (2016) and Cavaglia and Etheridge (2017) among others estimate the evolution of prices in abstract task-intensive and manual task-intensive compared to routine task-intensive occupation groups. Yet, one recurring caveat of such estimations is that the methods involved require potentially restrictive identification assumptions.

This paper proposes a new approach to estimate the changes of task prices under assumptions which are directly motivated by economic theory. I first show that, in the Roy (1951) model, workers' wage growth over time exclusively depends on the interplay between their task choices and these changing prices; and not on skill levels or any particular distribution of skills in the economy.[1] Intuitively, if a worker chooses a task for which the price is then raised in a comparative statics analysis, he will have an increased wage after the price change compared to a worker who chooses a task for which the (relative) price is lowered, even if one or both workers (endogenously) choose different tasks in the new equilibrium. This is because, from their revealed choices, we can infer that the former worker possesses skills that make him benefit from the new prices compared to the latter worker.

Second, I devise a method to empirically implement this insight in repeated cross-section data by noting that each individual in the population may be endowed with fundamental talents which fulfill two important conditions: their joint distribution among the workforce is stable over time ("comparability") and, for talent components that are observed in some datasets, they are significant predictors of workers' sorting into tasks ("first-stage"). Under these assumptions one can construct propensities to

[1] In the 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 efficiency unit of that task input (the "task price"). As in previous literature (e.g., Heckman and Sedlacek, 1985; Autor, Katz, and Kearney, 2006; Acemoglu and Autor, 2011; Firpo, Fortin, and Lemieux, 2013; Cortes, 2016), this paper assumes throughout that task prices change over time but not skills in tasks.

choose tasks conditional on talents and then, using the theoretical result, exploit wage growth over time conditional on the same talents to estimate the changes of task prices in the data.

I apply this new "Propensity Method" to the literature on routine-biased technical change (RBTC). Over the past decade, an active debate has developed about the question whether observed changes in the employment and wage structures are driven by the impact of RBTC on the labor market. In particular, the polarization of jobs away from routine and into abstract and manual tasks (e.g., Acemoglu and Autor, 2011; Goos, Manning, and Salomons, 2014) appears consistent with this theory. But the lack of polarization of the wage distribution in several countries (Dustmann, Ludsteck, and Schönberg, 2009; Naticchioni, Ragusa, and Massari, 2014; Green and Sand, 2015) and time periods (Mishel, Shierholz, and Schmitt, 2013) raises doubts about the importance of RBTC for wage inequality. One key implication of RBTC, which until recently has received less attention, is that task prices should polarize (i.e., abstract and manual prices should rise compared to routine).[2] I test this prediction by estimating changes of task prices in U.S. data.

To empirically 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 in a first-stage estimation and they arguably fulfill the comparability assumption.[3] Further, since the approach is derived from a model with discrete choices, 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 Occu-

[2]Simulations in the appendix demonstrate that different restrictions on the dependence structure of workers' skills can lead to wage distributions which are consistent with any of the previous findings. In contrast, RBTC's prediction that task prices should polarize is robust to these considerations.

[3]For example, the measures predict that, conditional on the other talents, an individual with high math ability is more likely to work in abstract tasks, an individual with high mechanical ability is likely to work mechanical tasks, and an individual with high verbal ability is more likely to either work in abstract or manual (i.e., largely service-oriented) tasks.

pational Information Network (O*NET). These broad occupation groups constitute the "tasks" for which prices are estimated.

The estimation results show that task prices indeed polarized 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. An individual-level prediction of RBTC, which is derived in the paper's appendix, is also borne out in the NLSY data: workers who according to their observable talents were more likely to work in rising non-routine tasks experienced higher wage growth than workers who were likely to work in declining routine tasks.[4] Finally, assigning workers the price changes of their chosen tasks in the initial period, plus an adjustment for the rising minimum wage, matches quite well the change of 27 year old males' wage inequality observed in U.S. data.

The general model that I set up in Section 2 allows for K different tasks and for amenities to affect task choices in addition to purely pecuniary considerations. I show that the theoretical intuition from above still applies but that wage growth of switchers also depends on the utility gains (or losses) from moving into tasks with different amenities. Empirically, the two considerations do not interfere, as changing task prices are identified from wage growth associated with individuals' *average* propensities in the two periods to work in different tasks whereas amenities are identified from *changing* propensities over time. I consider a version of the model with homogeneous amenity valuations across individuals and one with heterogeneous amenities (the latter leads to demanding requirements on the data). In the NLSY data, this estimation yields qualitatively similar task price changes as the pure Roy model while non-pecuniary benefits enjoyed in abstract and manual tasks are substantially higher than in routine tasks. This provides a new potential explanation for the fact that workers accept pay

[4]The main estimates 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 4.2 discusses the implications of RBTC versus competing or complementary hypotheses including international trade and offshoring, and SBTC combined with rising demand for services.

cuts when moving from higher-earning routine to lower-earning manual tasks (e.g., Cortes, 2016), and thus is again consistent with the observed job polarization predicted by RBTC.

The Propensity Method for 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.[5] Monte Carlo simulations show that the approach successfully recovers the actual changes in task prices under different assumptions about the dependence structure of workers' skills and, in the homogenous case, the average amenities across tasks. The Propensity Regression can also account for some additional confounding factors that may affect workers' wages, such as changing returns to college and the minimum wage. Section 3.5 provides a detailed comparison to recent alternative approaches for estimating task prices (e.g., Yamaguchi, 2012; Firpo, Fortin, and Lemieux, 2013; Gottschalk, Green, and Sand, 2015; Cortes, 2016; Yamaguchi, 2016).

The remainder of the paper proceeds as follows. The next section introduces the theoretical model and derives the relationship of worker sorting and wage growth with task price changes and amenities. Section 3 uses this relationship to establish the Propensity Method for estimating task prices. Then the economic hypotheses about RBTC, occupational grouping into tasks, and the main NLSY estimation sample are introduced. Section 5 presents the estimation results and analyzes their potential effect on the overall wage distribution. The last section concludes.

2 The Generalized Roy Model at the Individual Level

This section studies the relationship of changing task prices with workers' choices and wages in the generalized Roy model. A new equation linking individual workers' wage growth to task choices and non-pecuniary amenities is derived. In the appendix,

[5]Identification requires a local 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.

I show that other supposed effects of changing task prices on aggregate outcomes are not robust to different assumptions of how workers' skills are distributed across tasks.

2.1 Setup

Suppose there exist a discrete number of tasks $k = 1, ..., K$. In the empirical application, three tasks, abstract $k = A$, routine $k = R$, and manual $k = M$ will be used, but for now it is more convenient and more general to work with K tasks indexed by integers. Every worker i possesses a vector of log skills $\boldsymbol{s_i} = [s_{1i}\ s_{2i} \ldots s_{Ki}]'$ and faces log task prices $\boldsymbol{\pi_t} = [\pi_{1t}\ \pi_{2t} \ldots \pi_{Kt}]'$. Potential log wages are

$$w_{kit} = \pi_{kt} + s_{ki}\ \forall\ k \in \{1, ..., K\}. \tag{1}$$

Not giving s_{Ki} a time index t already imposes the key identification assumption detailed below, namely that (the conditional distributions of) workers' skills are not changing across cohorts whereas the task prices are changing over time.

In the generalized Roy model, workers have preferences over wages and non-pecuniary amenities of jobs. I will estimate different versions of this model, one with purely pecuniary preferences over tasks, one with homogeneous amenities for each task, which can then be priced, and the most general setup with heterogeneous non-pecuniary preferences across individuals. In the latter case, utility of worker i in task k at time t is:

$$U_{kit} = w_{kit} + v_{kit} \text{ with} \tag{2}$$

$$v_{kit} = a_{kt} + \boldsymbol{b_k'} X_i + e_{ki}. \tag{3}$$

Here a_{kt} is a non-pecuniary intercept in each task. The $\boldsymbol{b_k'} X_i$ are task-specific mappings from worker characteristics to utility. In practice, the valuation of each task k may differ by individuals' talent types. I further let idiosyncratic task valuations e_{ki} be mean zero and independent across individuals (they may be correlated across tasks for a given individual, however). Again, the key 'comparability' identification assumption is already imposed by making the (distribution of) idiosyncratic preferences e_{ki} and the heterogeneous mappings $\boldsymbol{b_k}$ time invariant. Finally, notice that only relative amenities in tasks, and thus the parameters a_{kt}, $\boldsymbol{b_k}$ compared to a chosen base group (the routine task in the empirical application), will be identifiable from workers' observed choices and wages.

Workers maximize their utility by choosing the task with the highest overall, pecuniary plus non-pecuniary, reward ($\boldsymbol{U_{it}} \equiv [U_{1it} \; U_{2it} \ldots U_{Kit}]'$):

$$U_{it} = w_{it} + v_{it} = \begin{cases} U_{1it} = w_{1it} + v_{1it} & \text{if } I_1(\boldsymbol{U_{it}}) = 1 \\ \quad \vdots & \\ U_{Kit} = w_{Kit} + v_{Kit} & \text{if } I_K(\boldsymbol{U_{it}}) = 1 \end{cases} \tag{4}$$

where $I_k(\boldsymbol{U_{it}}) \equiv \mathbf{1}[\max_{l=1,\ldots,K}\{U_{lit}\} = U_{kit}] = \mathbf{1}[U_{kit} \geq U_{lit} \; \forall \, l \neq k]$ is a choice indicator for task k.

The predominant interest of the following lies in estimating the changing task prices $\Delta\pi_k \equiv \pi_{k1} - \pi_{k0}$ between two points in time $t = 0$ and $t = 1$. In addition, I will derive how to estimate the average non-pecuniary valuations of tasks and, in Appendix B.2, the wage changes for workers who initially allocate to different tasks. Finally, the joint population distributions of workers' skills $F_{s_1,\ldots,s_K}$ (assumed not to change across cohorts, however) and idiosyncratic preferences $F_{e_1,\ldots,e_K}$ are left unrestricted.

2.2 The Effect of Changing Task Prices on Workers' Wages

Consider Equation (4) and the realized marginal utility change of worker i when potential utilities shift across tasks at time t:[6]

$$\mathrm{d}U_{it} = I_1(\boldsymbol{U_{it}})\,\mathrm{d}U_{1it} + \ldots + I_K(\boldsymbol{U_{it}})\,\mathrm{d}U_{Kit} = \sum_{k=1}^{K} I_k(U_{1it}, \ldots, U_{Kit})\,\mathrm{d}U_{kit}, \tag{5}$$

where d denotes a marginal change of the respective variable over time and the second equality writes out the dependency of choices on all K potential utilities. Due to the optimality of worker i's choice $I_k(\boldsymbol{U_{it}})$, a marginal change in task prices or non-pecuniary

[6] This may be because of potential wages (i.e., via task prices) or because of amenities changing or both.

components will only have a direct effect because workers' marginal utility changes are not affected by a potential reaction of their task choices. This is the envelope theorem, which implies that the effect of a marginal change is only the direct effect (those workers who switch tasks do not experience a utility gain from it on the margin).

I am however interested in discrete changes of potential utilities in general, and task prices in particular, between two points in time. Therefore, I integrate (5) from potential utilities $\{U_{1i0}, \ldots, U_{Ki0}\}$ to $\{U_{1i1}, \ldots, U_{Ki1}\}$. To get from marginal to discrete changes, hold constant $U_{li0}\ \forall\, l \geq 2$ first and integrate with respect to the potential utility in task 1:

$$U_{i|U_{1i1},U_{2i0},\ldots} - U_{i|U_{1i0},U_{2i0},\ldots} = \int_{U_{1i0}}^{U_{1i1}} I_1(U_{1it}, U_{2i0}, \ldots)\, \mathrm{d}U_{1it}.$$

Now integrate with respect to some U_{kit}, holding constant $U_{li0}\ \forall\, l > k$ at 0 as well as $U_{mi1}\ \forall\, m < k$ at 1. Then, $\forall\, k \in \{1, ..., K\}$:

$$U_{i|U_{1i1},\ldots,U_{ki1},\ldots,U_{Ki0}} - U_{i|U_{1i1},\ldots,U_{ki0},\ldots,U_{Ki0}} = \int_{U_{ki0}}^{U_{ki1}} I_k(U_{1i1}, ..., U_{kit}, ..., U_{Ki0})\, \mathrm{d}U_{kit}. \tag{6}$$

Summing all of these elements (6) from $k = 1$ to $k = K$ obtains

$$\begin{aligned} U_{i|U_{1i1},\ldots,U_{Ki1}} - U_{i|U_{1i0},\ldots,U_{Ki0}} = U_{i1} - U_{i0} = \Delta U_i = \\ = \Delta w_i + \Delta v_i = \sum_{k=1}^{K} \int_{U_{ki0}}^{U_{ki1}} I_k(U_{1i1}, ..., U_{kit}, ..., U_{Ki0})\, \mathrm{d}U_{kit}. \end{aligned} \tag{7}$$

Notice once again that the envelope theorem and the marginal result (5) are a tool to derive Equation (7), not an assumption. The assumption is the optimality of workers' choices in the sense that they maximize utility within this model. Appendix C shows a derivation of (7) with only two tasks that does not invoke the envelope theorem and where the complication of the different time indexes in $I_k(U_{1i1}, ..., U_{kit}, ..., U_{Ki0})$ does not appear in the main step of the derivation either.

Result (7) is rather intuitive: if a worker stays in his task k in both $t = 0$ and $t = 1$ ($I_k(\boldsymbol{U_{i1}}) = I_k(\boldsymbol{U_{i0}}) = 1$), his realized utility growth is equal to the change in his potential utility in the chosen task (i.e., $\Delta U_i = \Delta U_{ki}$). If the worker switches from one of the other tasks l to k, ($I_l(\boldsymbol{U_{i0}}) = 1$,$I_k(\boldsymbol{U_{i1}}) = 1$), he obtains part of the origin task's utility gain (or loss) as well as part of the destination task's utility gain, with the relative size of these parts determined by the points of indifference. Because of the different time indexes this looks complicated but it becomes clear when writing out Result (7) in a specific example:

$$\Delta U_i = \int_{U_{1i0}}^{U_{1i1}} I_1(U_{1it}, U_{2i0}, U_{3i0})\,\mathrm{d}U_{1it} + \int_{U_{2i0}}^{U_{2i1}} I_2(U_{1i1}, U_{2it}, U_{3i0})\,\mathrm{d}U_{2it} + \int_{U_{3i0}}^{U_{3i1}} I_3(U_{1i1}, U_{2i1}, U_{3it})\,\mathrm{d}U_{3it} \tag{8}$$

In this case with three tasks $k \in \{1, 2, 3\}$, suppose that the worker ranks the potential utilities as $U_{3i1} > U_{2i1} > U_{1i1} > U_{1i0} > U_{2i0} > U_{3i0}$. This implies that in the data he is observed switching from task 1 in $t = 0$ to task 3 in $t = 1$. Equation (8) then yields utility growth of $\Delta U_i = (U_{1i1} - U_{1i0}) + (U_{2i1} - U_{1i1}) + (U_{3i1} - U_{2i1}) = U_{3i1} - U_{1i0}$, which is exactly his realized utility in $t = 1$ minus $t = 0$. The second summand in (8) for example becomes $\int_{U_{1i1}}^{U_{2i1}} 1\,\mathrm{d}U_{2it} = U_{2i1} - U_{1i1}$, since at $U_{2it} = U_{1i1}$ the indicator $I_2(U_{1i1}, U_{2it}, U_{3i0})$ switches to 1.[7]

In order to estimate task prices and amenities, the aim is to derive a relationship between realized wages and task choices because these can be measured in the data. However, it is only possible to observe wages and choices at the end points of each period. In the case of the empirical application below, this means task choices and wages for the NLSY79 cohort ($t = 0$) and the NLSY97 cohort ($t = 1$), respectively. Therefore, I linearly interpolate the integrand of Equation (7)

$$I_k(U_{1i1}, ..., U_{kit}, ..., U_{Ki0}) \approx I_k(\boldsymbol{U_{i0}}) + \frac{I_k(\boldsymbol{U_{i1}}) - I_k(\boldsymbol{U_{i0}})}{U_{ki1} - U_{ki0}}(U_{kit} - U_{ki0}) \tag{9}$$

within $[I_k(\boldsymbol{U_{i0}}), I_k(\boldsymbol{U_{i1}})]$ to get

$$\Delta U_i = \Delta w_i + \Delta v_i \approx \sum_{k=1}^{K} \bar{I}_{ki} \Delta U_{ki} = \sum_{k=1}^{K} \bar{I}_{ki} \Delta w_{ki} + \sum_{k=1}^{K} \bar{I}_{ki} \Delta v_{ki}. \tag{10}$$

[7]Alternatively for a staying worker (in task 3, say), suppose he has utility ranking $U_{3i1} > U_{2i1} > U_{1i1} > U_{3i0} > U_{2i0} > U_{1i0}$. It looks like he is switching in each of the integrals, but in fact the utility growth implied by Equation (8) is (correctly) just the change of potential utility in the task that he is observed in at $t = 0$ and $t = 1$: $\Delta U_i = (U_{1i1} - U_{3i0}) + (U_{2i1} - U_{1i1}) + (U_{3i1} - U_{2i1}) = \Delta U_{3i}$.

Here $\bar{I}_{ki} \equiv \frac{I_k(\boldsymbol{U_{i1}})+I_k(\boldsymbol{U_{i0}})}{2}$ is the "average task choice" across the two periods and the last equality uses the definition of potential utility (2).

Proof. Replace the indicator $I_k(U_{1i1}, ..., U_{kit}, ..., U_{Ki0})$ for a specific k in Equation (7) with the linear interpolation:

$$\begin{aligned}\int_{U_{ki0}}^{U_{ki1}} I_k(U_{1i1}, ..., U_{kit}, ..., U_{Ki0})\,\mathrm{d}U_{kit} &\approx \int_{U_{ki0}}^{U_{ki1}} \left[I_k(\boldsymbol{U_{i0}}) + \frac{I_k(\boldsymbol{U_{i1}}) - I_k(\boldsymbol{U_{i0}})}{U_{ki1} - U_{ki0}} (U_{kit} - U_{ki0}) \right] \mathrm{d}U_{kit} \\ &= I_k(\boldsymbol{U_{i0}})\Delta U_{ki} + \frac{I_k(\boldsymbol{U_{i1}}) - I_k(\boldsymbol{U_{i0}})}{U_{ki1} - U_{ki0}} \left[\frac{1}{2} U_{kit}^2 - U_{ki0} U_{kit} \right]_{U_{ki0}}^{U_{ki1}} \\ &= I_k(\boldsymbol{U_{i0}})\Delta U_{ki} + \frac{1}{2}(I_k(\boldsymbol{U_{i1}}) - I_k(\boldsymbol{U_{i0}}))(U_{ki1} - U_{ki0}) \\ &= \bar{I}_{ki}\Delta U_{ki}\end{aligned}$$

Summing this up over all k gives Equation (10). □

The intuition in Equation (10) remains the same as before: if a worker stays in his task, his realized utility gain is the change of his potential utility in that task. That is, $\Delta U_i = \Delta U_{ki}$ if $I_k(\boldsymbol{U_{i1}}) = I_k(\boldsymbol{U_{i0}}) = 1$, which is actually not an approximation. If the worker switches (e.g., tasks l to k, $I_l(\boldsymbol{U_{i0}}) = 1$, $I_k(\boldsymbol{U_{i1}}) = 1$), he obtains part of the origin task's utility gain (or loss) as well as part of the destination task's utility gain, set to exactly half-half by the interpolation (i.e., $\Delta U_i = \frac{1}{2}\Delta U_{li} + \frac{1}{2}\Delta U_{ki}$). For an individual worker this may seem quite off, as indifference might occur close to the initial task price/amenity vector and the gain is mostly from the destination task's changing utility, or it might occur close to the final prices/amenities and the gain is mostly from the origin task. Nonetheless, when I take expectations below over all workers with observable talents $\boldsymbol{x}$, it will turn out a good approximation of the average utility (and wage) gain for a given switch.[8]

Notice also that in the interpolation (9) I employ $I_k(\boldsymbol{U_{i0}}) = I_k(U_{1i0}, ..., U_{ki0}, ..., U_{Ki0})$ as well as $I_k(\boldsymbol{U_{i1}}) = I_k(U_{1i1}, ..., U_{ki1}, ..., U_{Ki1})$ instead of $I_k(U_{1i1}, ..., U_{ki0}, ..., U_{Ki0})$ as well as $I_k(U_{1i1}, ..., U_{ki1}, ..., U_{Ki0})$. The reason is that the former are observed choices in the

[8]See the discussion of the Monte Carlo simulations in Section E. Also notice that the utility gain (and thus the approximation error) is bounded by the utility changes of the two tasks that a switching worker chooses. That is, if he moves from l to k, we have $\Delta U_i \in [\Delta U_{li}, \Delta U_{ki}]$.

data, and can therefore be used in the empirics, but also that the latter may lead to nonsensical results.[9] The Monte Carlo simulations in Section E indicate that also this choice is innocuous for identifying the correct task prices.

To further simplify Result (10), consider $v_{it} = \sum_k I_k(\boldsymbol{U_{it}})v_{kit}$ and then write

$$\begin{aligned}\Delta v_i &= \sum_k I_k(\boldsymbol{U_{i1}})v_{ki1} - \sum_k I_k(\boldsymbol{U_{i0}})v_{ki0} \\ &= \sum_k \bar{I}_{ki}\Delta v_{ki} + \sum_k \bar{v}_{ki}\Delta I_{ki},\end{aligned}$$

with $\bar{v}_{ki} \equiv \frac{1}{2}(v_{ki0} + v_{ki1}) = \bar{a}_k + \boldsymbol{b_k'}X_i + e_{ki}$ and $\Delta I_{ki} \equiv I_k(\boldsymbol{U_{i1}}) - I_k(\boldsymbol{U_{i0}})$. Inserting this into Equation (10), the realized wage growth of individual worker i in the generalized Roy model becomes:

$$\Delta w_i \approx \sum_k \bar{I}_{ki}\Delta\pi_k - \sum_k \bar{v}_{ki}\Delta I_{ki}, \tag{11}$$

where I have now substituted $\Delta w_{ki} = \Delta\pi_k$, since I assume s_{ki} to be time-invariant throughout this paper. That is, under the identification Assumptions 1 and 2 below, I hold the conditional (fundamental) skill distribution in the population constant between time $t = 0$ and $t = 1$.

Equation (11) has firstly a purely pecuniary part ($\sum_k \bar{I}_{ki}\Delta\pi_k$): if a worker stays in his task, his wage gain is the change in the price of that task. If the worker switches, he obtains half of the origin task's price change as well as half of the destination task's price change. The strength of this result is that it accommodates endogenous switches, i.e., which are due to changes in task prices or amenities. Provided that amenities are

[9] For example, when using $I_k(U_{1i1}, ..., U_{ki0}, ..., U_{Ki0})$ and $I_k(U_{1i1}, ..., U_{ki1}, ..., U_{Ki0})$ in the three-task Equation (8), the approximation yields:

$$\begin{aligned}\Delta U_i \approx &\frac{1}{2}\left[I_1(U_{1i0}, U_{2i0}, U_{3i0}) + I_1(U_{1i1}, U_{2i0}, U_{3i0})\right]\Delta U_{1i} \\ &+\frac{1}{2}\left[I_2(U_{1i1}, U_{2i0}, U_{3i0}) + I_2(U_{1i1}, U_{2i1}, U_{3i0})\right]\Delta U_{2i} \\ &+\frac{1}{2}\left[I_3(U_{1i1}, U_{2i1}, U_{3i0}) + I_3(U_{1i1}, U_{2i1}, U_{3i1})\right]\Delta U_{3i}\end{aligned}$$

Now suppose $U_{3i1} > U_{2i1} > U_{1i1} > U_{3i0} > U_{2i0} > U_{1i0}$. This gives $\Delta U_i \approx \frac{1}{2}\Delta U_{1i} + \frac{1}{2}\Delta U_{2i} + \frac{1}{2}\Delta U_{3i}$, which is quite off even for this worker who stays in task 3. Instead, the approximation (9) and the resulting Equation (10) yield exactly the correct utility growth $\Delta U_i \approx \Delta U_{3i}$.

controlled for in an adequate manner, the first term of (11) suggests that task prices can be recovered from a regression of first-differenced wages on "average" task choices $\bar{I}_{ki}$, which are straightforwardly constructed using the two cohorts of the NLSY.

The second summand on the right of Equation (11) is then the intuitive extension of a purely pecuniary model: with optimal choices, a worker's observed wage growth is the change in the prices of his chosen tasks minus the utility gain (loss) from the behavioral response of switching tasks. That is, if a utility-optimizing worker chooses to switch tasks (e.g., from l to k so that $\Delta I_{ki} = 1$ and $\Delta I_{li} = -1$), we observe lower wage growth than the change in relevant task prices when he gains amenities (i.e., $\bar{v}_{ki} > \bar{v}_{li}$ and thus $\sum_k \bar{v}_{ki}\Delta I_{ki} > 0$) via the move. Vice versa, we observe higher wage growth than the task price changes when he moves to a less desirable task ($\sum_k \bar{v}_{ki}\Delta I_{ki} < 0$).

Notice in Equation (11) it is the *average* amenity over both periods $\bar{v}_{ki}$ that the worker is moving into which matters for wage changes. For a switcher from l to k,

$$\sum_k \bar{v}_{ki}\Delta I_{ki} = \bar{v}_{ki} - \bar{v}_{li} = \frac{1}{2}(v_{ki1} - v_{li1}) + \frac{1}{2}(v_{ki0} - v_{li0}),$$

conditional on wage gains associated with average choices $\sum_k \bar{I}_{ki}\Delta\pi_k$, moving into the currently high-amenity task (i.e., $v_{ki1} - v_{li1} > 0$) is offset with lower wage growth. But also moving into a task that last period carried high amenities ($v_{ki0} - v_{li0} > 0$) is associated with lower wage growth because it implies that the worker was compensated last period for working in the low-amenity task, which now falls away with the switch. Both of these factors enter equally into the wage equation (11). Hence one cannot distinguish them empirically and only identify the average amenity over the two periods.

Finally, notice when the worker makes no switch, the amenity considerations do not come into play at all (i.e., $\sum_k \bar{v}_{ki}\Delta I_{ki} = 0$) and the changing wage is just the changing task price.

3 The Propensity Method for Estimating Task Prices and Amenities

Result (11) 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. However, among other challenges, such an approach would critically rely on disentangling task price changes over time from individuals' skill accumulation over the life-cycle, which occurs even in the absence of any other changes. This would involve modeling the process of how skills evolve systematically by past and current task choices, age, and other variables as well as allowing for idiosyncratic shocks to the skills or due to employer learning. Panel estimation would therefore be difficult to do convincingly (at least without high-quality longitudinal information for several cohorts of workers; also see the discussion in Section 3.5). 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 (and potentially their non-pecuniary preferences) depend on these talents in different ways.

3.1 Identification Assumptions

In order to discuss the empirical identification assumptions, I start with the pure Roy model without non-pecuniary amenities. That is, $v_{kit} = 0 \; \forall \, k, i, t$ and therefore utility $U_{kit} = w_{kit} = \pi_{kt} + s_{ki}$ only depends on skills and prices now. Equation (11) becomes

$$w_{i1} - w_{i0} \approx \sum_k \bar{I}_{ki} \Delta\pi_k = \sum_k \frac{I_k(\boldsymbol{s_i}, \boldsymbol{\pi_1}) + I_k(\boldsymbol{s_i}, \boldsymbol{\pi_0})}{2} \Delta\pi_k, \tag{12}$$

where the dependence of choices on skills and task prices is made explicit in the rightmost expression. I also consider a specific functional form for the skills. In particular, a natural formulation is Heckman and Sedlacek (1985)'s linear factor model of log wages:

$$w_{kit} = \pi_{kt} + s_{ki} = \pi_{kt} + \boldsymbol{\beta}_k' \boldsymbol{x_i} + u_{ki}, \tag{13}$$

where $\boldsymbol{x_i} = [x_{1i}\ x_{2i} \ldots x_{Ji}]'$ are a vector of observed talents, $\boldsymbol{\beta_k}$ the corresponding linear projection coefficients, and $u_{ki} = \boldsymbol{\delta_k}'\boldsymbol{z_i}$ a regression error, which again depends on unobservable talents $\boldsymbol{z_i}$ and linear projection coefficients $\boldsymbol{\delta_k}$. 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. As in Heckman and Sedlacek (1985), Firpo, Fortin, and Lemieux (2013), and throughout the RBTC literature, the $\boldsymbol{\beta_k}$ and $\boldsymbol{\delta_k}$ vectors are assumed task-specific but time-invariant, while the task prices π_{kt} are changing with RBTC over time (see discussion in Section 4). Although all the results in this paper hold for a general time-invariant dependency of s_{ki} on observables $\boldsymbol{x_i}$ and unobservables $\boldsymbol{z_i}$, specification (13) is adopted for better illustration from now on.

Empirically, the estimation approach for task prices (and amenities) proposed below requires data on worker talents $\boldsymbol{x}$ that fulfill the following two assumptions:

Assumption 1. *[First-stage] The vector $\boldsymbol{x}$ predicts a worker i's task choices in both periods of time $t \in \{0, 1\}$.*

Assumption 2. *[Comparability] Individuals with the same $\boldsymbol{x}$ vector are comparable over time. That is, conditional on $\boldsymbol{x}$ for all $\boldsymbol{x}$, the population distributions of unobservable talents $\boldsymbol{z}$ are the same in both periods.*[10]

Let $C_i \in \{0, 1\}$ denote an indicator for whether individual i is from the initial cohort at $t = 0$ (NLSY79 in the empirical application) or the later cohort at $t = 1$ (NLSY97). In the data, one can always compute conditional expected choices $\mathrm{E}(I_k(\mathbf{s_i}, \boldsymbol{\pi_t}) \mid \mathbf{x_i} = \mathbf{x}, C_i = t)$ and wages $\mathrm{E}(w_{it} \mid \mathbf{x_i} = \mathbf{x}, C_i = t)$ in each respective cohort. What Assumption 2 then does is to ensure that these expectations do not in fact depend on C_i. That is, the actual and the counterfactual conditional on $\boldsymbol{x}$ are the same in both cohorts:

$$\mathrm{E}(w_{it} \mid \mathbf{x_i} = \mathbf{x}, C_i = 1) = \mathrm{E}(w_{it} \mid \mathbf{x_i} = \mathbf{x}, C_i = 0) = \mathrm{E}(w_{it} \mid \mathbf{x_i} = \mathbf{x}) \tag{14}$$

$$\mathrm{E}(I_k(\mathbf{s_i}, \boldsymbol{\pi_t}) \mid \mathbf{x_i} = \mathbf{x}, C_i = 1) = \mathrm{E}(I_k(\mathbf{s_i}, \boldsymbol{\pi_t}) \mid \mathbf{x_i} = \mathbf{x}, C_i = 0) = \mathrm{E}(I_k(\mathbf{s_i}, \boldsymbol{\pi_t}) \mid \mathbf{x_i} = \mathbf{x}) \tag{15}$$

[10] In the case of heterogeneous amenities below, also the population distributions of idiosyncratic preferences $\{\boldsymbol{e_k}\}$ are the same in both periods.

Therefore, I can take expectations on both sides of (12) conditional on worker i's vector of talents $\boldsymbol{x}$ only to get:[11]

$$\mathrm{E}(w_{i1} \mid \boldsymbol{x_i} = \boldsymbol{x}) - \mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) \approx \sum_k \frac{p_k(\boldsymbol{x}, \boldsymbol{\pi_1}) + p_k(\boldsymbol{x}, \boldsymbol{\pi_0})}{2} \Delta\pi_k, \tag{16}$$

where

$$p_k(\boldsymbol{x}, \boldsymbol{\pi_t}) \equiv \mathrm{E}[I_k(\boldsymbol{s_i}, \boldsymbol{\pi_t}) \mid \boldsymbol{x_i} = \boldsymbol{x}] = \Pr[-(u_{ki} - u_{li}) \leq \pi_{kt} - \pi_{lt} + (\boldsymbol{\beta}_k - \boldsymbol{\beta}_l)'\boldsymbol{x} \; \forall \, l \neq k] \tag{17}$$

Given the availability of talent measures $\boldsymbol{x}$, Equation (16) can be empirically implemented by computing outcomes $\mathrm{E}(w_{it} \mid \boldsymbol{x_i} = \boldsymbol{x})$ and explanatory variables $p_k(\boldsymbol{x}, \boldsymbol{\pi_t})$ in each period. Details of the implementation are explained in the next section. What is necessary is that this strategy is not confounded by unobserved selection effects. In particular, it should be the same whether one computes $\mathrm{E}(w_{i1} \mid \mathbf{x_i} = \mathbf{x})$ and $p_k(\boldsymbol{x}, \boldsymbol{\pi_1})$ using measurements of $\boldsymbol{x}$ from period $t = 1$ (actual) or $t = 0$ (counterfactual), and accordingly for $p_k(\boldsymbol{x}, \boldsymbol{\pi_0})$. This is what Assumption 2 and its implications (14)-(15) ensure.[12]

The NLSY dataset used in the empirical application below is attractive in terms of comparability because it provides pre-labor market talents $\boldsymbol{x}$ 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),

[11] Without Assumption 2,

$$\mathrm{E}(w_{i1} \mid \mathbf{x_i} = \mathbf{x}, C_i = 0) - \mathrm{E}(w_{i0} \mid \mathbf{x_i} = \mathbf{x}, C_i = 0) \approx$$

$$\approx \sum_k \frac{\mathrm{E}(I_k(\mathbf{s_i}, \boldsymbol{\pi_1}) \mid \mathbf{x_i} = \mathbf{x}, C_i = 0) + \mathrm{E}(I_k(\mathbf{s_i}, \boldsymbol{\pi_0}) \mid \mathbf{x_i} = \mathbf{x}, C_i = 0)}{2} \Delta\pi_k,$$

(or conditioning everywhere on $C_i = 1$ instead) would be mathematically correct but not implementable in the observed data.

[12] Comparability achieves this as it assumes that the unobservable skill selection into observable groups $\boldsymbol{x}$ does not change over time. Thus, individuals with the same $\boldsymbol{x}$ vector are "in distribution" equally skilled in both periods ($F_{s_1,\ldots,s_K|\boldsymbol{x}}$ does not carry a time index).

comparability should hold in that data. Details of robustness checks in this regard and control variables are discussed below.

The first-stage Assumption 1 identifies the parameters in the sense that the estimation model can be solved for the unknown variables. Suppose in contrast that talents do not discriminate between tasks, i.e., that $\bar{p}_k(\boldsymbol{x}) \equiv \frac{p_k(\boldsymbol{x},\boldsymbol{\pi_1})+p_k(\boldsymbol{x},\boldsymbol{\pi_0})}{2} = \bar{p}_k$ is independent of $\boldsymbol{x}$. In this case, (16) becomes one equation with K unkowns (the $\Delta\pi_k$s). Therefore, in order to identify all (relative) task prices, $\boldsymbol{x}$ has to predict all of them. That is, if it cannot discriminate between two tasks, one cannot identify their (relative) price changes.[13]

The purpose of the first-stage is thus to provide good predictors of task choice probabilities conditional on $\boldsymbol{x}$ and $\boldsymbol{\pi_t}$. 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 $\boldsymbol{x}$ realization.[14] The requirement of Assumption 1 that $\boldsymbol{x}$ predict task choices *in both periods of time* becomes critical when I add non-pecuniary amenities below.

Economically, Assumption 1 demands that for K tasks there exist at least $K-1$ elements of the $\boldsymbol{x}$ vector that generate qualitatively independent task predictions. In the empirical application using NLSY data below, conditional on the other talents, the math talent predicts the abstract task (the element corresponding to math in $\boldsymbol{\beta}_A$ is high; compare Equation (17)) and the mechanical talent predicts the routine task (the element corresponding to mechanical in $\boldsymbol{\beta}_R$ 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).

[13] To see this best, consider three tasks as in the application below, and suppose that $\boldsymbol{x}$ does not discriminate between task 1 and 2. Rewriting (16) relative to the first task price ($\bar{p}_1 + \bar{p}_2 + \bar{p}_3(\boldsymbol{x}) = 1$) gives

$$\mathrm{E}(w_{i1} \mid \boldsymbol{x_i} = \boldsymbol{x}) - \mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) \approx \Delta\pi_1 + \bar{p}_2 \cdot (\Delta\pi_2 - \Delta\pi_1) + \bar{p}_3(\boldsymbol{x}) \cdot (\Delta\pi_3 - \Delta\pi_1).$$

In this case, the price change of task 3 relative to 1 can be identified from variation across individuals with different $\boldsymbol{x}$. However, the other prices cannot be identified, as infinitely many combinations of $\Delta\pi_1$ and $\Delta\pi_2$ solve the model.

[14] 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) and analyzed in an earlier version of this paper (Böhm, 2017), the actual choice frequencies for discrete demographic $\boldsymbol{x}$ cells were 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 (though they do of course sum to 1).

3.2 Empirical Implementation

The parameters in Equation (16) can be estimated in different ways. The easiest is to note that (16) can be rewritten as a regression equation:

$$w_{it} \approx \mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) + \sum_k \Delta\pi_k \cdot \overline{p}_k(\boldsymbol{x_i}) \cdot \mathbf{1}[t=1] + e_{it},$$

where $w_{it} = \mathrm{E}(w_{it} \mid \boldsymbol{x_i} = \boldsymbol{x}) + e_{it}$ and $\mathbf{1}[t=1]$ indicates a dummy for time period $t=1$. Notice that e_{it} is mean zero as well as uncorrelated with all the elements of $\boldsymbol{x_i}$ and hence also uncorrelated with $\overline{p}_k(\boldsymbol{x_i})\ \forall\, k$. Now, in order to identify the correct $\Delta\pi_k$, I need to model $\mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) = f(\boldsymbol{x_i})$ such that the error term $\varepsilon_{it} = e_{it} + \mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) - f(\boldsymbol{x_i})$ in the estimable regression equation is uncorrelated with the regressors $\overline{p}_k(\boldsymbol{x_i}) \cdot \mathbf{1}[t=1]$. Specifying $f(\boldsymbol{x_i}) = \sum_{k=1}^K \theta_k \cdot \overline{p}_k(\boldsymbol{x_i})$ ensures this, since the regression residual $\mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) - \sum_{k=1}^K \hat{\theta}_k \cdot \overline{p}_k(\boldsymbol{x_i})$ is by construction uncorrelated with $\overline{p}_k(\boldsymbol{x_i})\ \forall\, k$. In addition, one could include meaningful further regressors into $f(\boldsymbol{x_i})$ on top of $\sum_{k=1}^K \theta_k \cdot \overline{p}_k(\boldsymbol{x_i})$. This may potentially improve the fit of regression (18), and thus the standard errors of the estimated $\hat{\Delta\pi}_k$, but it should not change the probability limit of $\hat{\Delta\pi}_k$. As a check, I will report one richer such specifications in each of the results tables below.

However, the "main" regression implementation of Equation (16) using individual-level wages becomes:

$$w_{it} = \sum_{k=1}^K \theta_k \cdot \overline{p}_k(\boldsymbol{x_i}) + \sum_{k=1}^K \Delta\pi_k \cdot \overline{p}_k(\boldsymbol{x_i}) \cdot \mathbf{1}[t=1] + \varepsilon_{it}, \tag{18}$$

with propensities $\overline{p}_k(\boldsymbol{x_i}) \equiv \frac{p_k(x_i,\pi_1)+p_k(x_i,\pi_0)}{2}$. Intuitively, in Equation (18) I want to estimate the *change of* the return to $\overline{p}_k(\boldsymbol{x_i})$ between periods $t=0$ and $t=1$, which gives the wage gain from comparative advantage in task k over time. For this, I need to fix the level of wages for $\mathrm{E}(w_{i0} \mid \{\overline{p}_k(\boldsymbol{x_i})\}_k)$ so that it is not confounded by baseline different wages for workers of different choices $\overline{p}_k(\boldsymbol{x_i})$. The most parsimonious specification that ensures this is the linear $\mathrm{E}(w_{i0} \mid \{\overline{p}_k(\boldsymbol{x_i})\}_k) = \sum_{k=1}^K \theta_k \cdot \overline{p}_k(\boldsymbol{x_i})$. Pooling the data over two periods, and using time interactions for the regressors of interest $\overline{p}_k(\boldsymbol{x_i})$, therefore allows to conveniently estimate the correct $\Delta\pi_k$s. The Monte Carlo simulations in

Section E confirm that this "Propensity Regression" approach works.[15]

The Monte Carlo simulations in the Appendix examine to what extent the Propensity Regression (18) can allow for confounding factors that affect wages. Broadly, some types of changing skill accumulation (e.g., college attainment or labor market experience) or changing returns to skills (returns to college or experience) can be accounted for separately, but only in some cases 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 of accumulated skill via (task-specific) labor market experience is unlikely to be very strong at age 27.[16] Also, what the Monte Carlos suggest is important is to run different specifications (i.e., controlling for changing returns to college or education more broadly, or adjusting for the minimum wage) and to check whether differences in the parameter estimates arise. I do this in the empirical application below.

One alternative approach to identifying the task prices is based on minimum distance estimation of Equation (16). Briefly, the idea is that Equation (12) holds for every element of the $\boldsymbol{x}$ vector, thereby delivering J different moment conditions when $\boldsymbol{x} = [x_1 \dots x_j \dots x_J]'$. Estimation then relates the moment conditions to each other in a quadratic form. The detailed procedure is explained in the Appendix, and its estimates are reported as an alternative in Tables 2 and 3 below. While less straightforward than the Propensity Regression, one attractive feature of the minimum distance is that it also provides an over-identification test ("J-test") of the model, because Equation (12) will naturally not be exactly correct for every element x_j in the data.

[15]Notice that there is no additional assumption made regarding ε_{it}. This regression error simply consists of the individual i's deviation from the conditional wage $w_{it} - \mathrm{E}(w_{it} \mid \boldsymbol{x_i} = \boldsymbol{x})$ plus the difference $\mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) - \sum_{k=1}^{K} \theta_k \cdot \overline{p}_k(\boldsymbol{x_i})$. What is important is that these are both—by definition and mechanically, respectively—uncorrelated with $\overline{p}_k(\boldsymbol{x_i})$.

[16]For example, Ashworth, Hotz, Maurel, and Ransom (2017) document that early labor market experience in the NLSY97 is overall similar to the NLSY79 but the composition did shift from out-of-school to in-school work experience, which may to some extent have changed the task-specific experiences as well. Another advantage of using young workers, according to Gottschalk, Green, and Sand (2015), is that their wages may better capture recent task price movements compared to older workers, whose wages are more sticky (e.g., because of implicit contracting).

3.3 Homogenous Amenities Across Individuals

I now reintroduce the non-pecuniary amenities across tasks. An important benchmark case is that of homogenous amenities, that is, the tasks carry different non-pecuniary values but they benefit every worker to the same extent (one could think of working hours, health benefits, and so on). This is for example consistent with workers' preferences over occupations in Lee and Wolpin (2006) or Keane and Wolpin (1997).

With homogenous amenities, Equation (3) effectively becomes $v_{kit} = v_{kt} = a_{kt}$ and $\bar{v}_{ki} = \bar{a}_k = \frac{1}{2}(a_{k0} + a_{k1})$ because non-task-specific $\boldsymbol{b}_{\boldsymbol{t}}' X_i + e_i$ affect neither choices nor wages. In this case, taking expectations conditional on $\boldsymbol{x_i} = \boldsymbol{x}$ on both sides of (11) gives

$$\begin{aligned} \mathrm{E}(w_{i1} \mid \boldsymbol{x_i} = \boldsymbol{x}) - \mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) &\approx \sum_k \mathrm{E}[\bar{I}_{ki} \Delta \pi_k \mid \boldsymbol{x}] - \sum_k \mathrm{E}[\bar{a}_k \Delta I_{ki} \mid \boldsymbol{x}] \\ &= \sum_k \bar{p}_k(\boldsymbol{x}) \Delta \pi_k - \sum_k \Delta p_k(\boldsymbol{x}) \bar{a}_k, \end{aligned} \tag{19}$$

where I could factor the constants $\Delta \pi_k$ and $\bar{a}_k$ out of the expectation and then define $\bar{p}_k(\boldsymbol{x}) \equiv \frac{p_k(\boldsymbol{x}, \boldsymbol{\pi_1}, \boldsymbol{a_1}) + p_k(\boldsymbol{x}, \boldsymbol{\pi_0}, \boldsymbol{a_0})}{2}$ and $\Delta p_K(\boldsymbol{x}) \equiv p_K(\boldsymbol{x}, \boldsymbol{\pi_1}, \boldsymbol{a_1}) - p_K(\boldsymbol{x}, \boldsymbol{\pi_0}, \boldsymbol{a_0})$. Notice that, as before, even with homogenous amenities, task choices $I_k(\boldsymbol{x}, \boldsymbol{\pi_t}, \boldsymbol{a_t})$ differ across individuals because of differences in observable ($\boldsymbol{x}$) and unobservable ($\boldsymbol{z}$) talents.[17]

Equation (19) is the key result for estimating task prices in the presence of amenities. The first summand on the bottom right hand side implies as before that "workers of type" $\boldsymbol{x}$ experience wage growth according to the price changes of the tasks that they on average sort into. The second summand implies that "excess" rising wages, i.e., conditional on average task sorting $\bar{p}_k(\boldsymbol{x})$, that are associated with flows of $\boldsymbol{x}$ into k ($\Delta p_k(\boldsymbol{x}) > 0$) indicate falling amenities for this type of workers, and vice versa. Therefore, workers' behavioral response of moving into valuable amenities has to be subtracted from their wage gain. Those worker types $\boldsymbol{x}$ that move into amenity-rich tasks (with high $\bar{a}_k$) have lower wage gains, ceteris paribus. Empirically, Equation (19) can be used to employ a Propensity Regression approach as above, with the addition

[17] The choice propensity (17) becomes $p_k(\boldsymbol{x}, \boldsymbol{\pi_t}, \boldsymbol{a_t}) \equiv \mathrm{E}[I_k(\boldsymbol{s_i}, \boldsymbol{\pi_t}, \boldsymbol{a_t}) \mid \boldsymbol{x_i} = \boldsymbol{x}] = \Pr[\max_{l=1,\ldots,K}\{U_{lit}\} = U_{kit} = \pi_{kt} + \boldsymbol{\beta}_{\boldsymbol{k}}' \boldsymbol{x} + u_{ki} + a_{kt}]$.

that one needs to account for the utility gain from the behavioral response. That is, I need to control for the *change of* sorting $\Delta p_K(\boldsymbol{x})$ in the estimation:

$$
\begin{aligned}
w_{it} = & \sum_{k=1}^{K} \theta_k \cdot \overline{p}_k(\boldsymbol{x_i}) + \sum_{k=1}^{K} \psi_k \cdot \Delta p_k(\boldsymbol{x_i}) \\
& + \sum_{k=1}^{K} \Delta\pi_k \cdot \overline{p}_k(\boldsymbol{x_i}) \cdot \mathbf{1}[t=1] + \sum_{k=1}^{K} (-\bar{a}_k) \cdot \Delta p_k(\boldsymbol{x_i}) \cdot \mathbf{1}[t=1] + \varepsilon_{it}
\end{aligned} \tag{20}
$$

One attractive feature of the approach at hand is that the *average* amenities are also identified, just as the *changing* task prices. In particular, the average amenity in each task is simply the estimation coefficient (-$\hat{\bar{a}}_k$) on the change in sorting regressor $\Delta p_k(\boldsymbol{x})$. Therefore, the approach readily delivers a valuation of the different tasks in dollar terms, that is, it prices the amenities. This is conceptually different from recent papers, which use the size of workers' flows across firms (Sorkin, 2018) or occupations (Cortes and Gallipoli, 2017) in a separate analysis from the wages in order to evaluate firm amenities and the (pecuniary and non-pecuniary) costs of occupational mobility, respectively. Here, we instead exploit the systematic interaction between moves (propensity changes) and wage changes, which may be considered a more direct way of measuring amenity values. The approach requires distinguishing between average propensities (i.e., worker type $\boldsymbol{x}$'s comparative advantage) and changing propensities ($\boldsymbol{x}$'s shift in amenities) for identification.

Finally, one could have had the ex ante intuition that the Propensity Method would break down when amenities are not time-constant, since one would not be able to disentangle the contribution to realized wage growth of (changing) sorting due to changing task prices and changing task amenities. However, as discussed in relation to Equation (11), moving into current as well as past amenities both matter equally for wage changes. Therefore, it turns out that this contribution to wage growth is only via the *changing* sorting $\Delta p_k(\boldsymbol{x})$ into *average* amenities, whereas the contribution to wage growth from *changing* task prices is only via the *average* sorting $\bar{p}_k(\boldsymbol{x})$. The Propensity Method is thus robust to both changing task prices and changing amenities over time.

The Appendix reports on Monte Carlo experiments using Propensity Regression (20) in the presence of homogenous amenities. It turns out that, indeed, one can estimate the task prices consistently and precisely when including the controls for worker type $\boldsymbol{x}$'s change in sorting $\sum_k(-\bar{a}_k) \cdot \Delta p_k(\boldsymbol{x})$. Under reasonable parameter values, task prices from the original Propensity Method without this control are biased, but moderately so. An attractive feature of this estimation is that one can also identify the $\bar{a}_k$ (minus one omitted task), that is, the relative amenity of the different tasks valued in dollar terms.

3.4 Heterogeneous Tastes for Amenities

The general case with heterogeneous tastes for amenities across workers is somewhat more difficult. Considering Equations (3) and (11) again, I am interested in the average amenity across the two periods: $\bar{v}_{ki} = \bar{a}_k + \boldsymbol{b}_{\boldsymbol{k}}' X_i + e_{ki}$. I assume in the following that the talents $\boldsymbol{x_i}$, which enter workers' skills, are the same as the X_i that load on the non-pecuniary valuations via $\boldsymbol{b_k}$. As we will see, the task prices are still identified without an exclusion restriction on X_i.

Taking expectations conditional on $\boldsymbol{x_i} = \boldsymbol{x}$, extended Equation (11) becomes

$$\begin{aligned} \mathrm{E}(w_{i1} \mid \boldsymbol{x_i} = \boldsymbol{x}) - \mathrm{E}(w_{i0} \mid \boldsymbol{x_i} = \boldsymbol{x}) \approx & \sum_k \bar{p}_k(\boldsymbol{x}) \Delta \pi_k - \sum_k \Delta p_k(\boldsymbol{x}) \bar{a}_k \\ & - \sum_k \boldsymbol{b}_{\boldsymbol{k}}' \boldsymbol{x} \Delta p_k(\boldsymbol{x}) - \sum_k \mathrm{E}[e_{ki} \Delta I_{ki} \mid \boldsymbol{x}], \end{aligned} \tag{21}$$

with the top two terms the same as in Equation (19). I now have $\sum_k \mathrm{E}[\boldsymbol{b}_{\boldsymbol{k}}' \boldsymbol{x} \Delta I_{ki} \mid \boldsymbol{x}] = \sum_k \boldsymbol{b}_{\boldsymbol{k}}' \boldsymbol{x} \cdot \Delta p_k(\boldsymbol{x})$ in the bottom, which constitutes differences by $\boldsymbol{x}$-types of how changing sorting into task k across the two periods $t = 0$ and $t = 1$ is valued in terms of amenities. Again, this can in principle be controlled for in Propensity Regressions by introducing an interaction term of the observable $\boldsymbol{x}$-vector with the estimated sorting changes $\Delta p_k(\boldsymbol{x})$. Since the latter is also based on $\boldsymbol{x}$ as well as enters as a separate regressor in levels, however, this may in some cases lead to an (imperfect) multicollinearity problem.

The second term in the bottom of Equation (21) is due to the worker's idiosyncratic valuation of different tasks. I interpret potential amenity Equation (3) as a linear projection, which renders e_{ki} a regression error that is uncorrelated with $\boldsymbol{x}$. Nonetheless, one cannot factor e_{ki} out of the conditional expectation because it is a partial derminant of and thus correlated with ΔI_{ki}, which in turn is correlated with (partially determined by) $\boldsymbol{x}$.[18] Instead, I can write the conditional expectation as $\mathrm{E}[e_{ki}\Delta I_{ki} \mid \boldsymbol{x}] = cov[e_{ki}\Delta I_{ki} \mid \boldsymbol{x}]$ or

$$\mathrm{E}[e_{ki}\Delta I_{ki} \mid \boldsymbol{x}] = \mathrm{E}\left[e_{ki}\mathrm{E}[\Delta I_{ki} \mid \boldsymbol{x}, e_{ki}]\right] = \mathrm{E}[e_{ki}\Delta p_k(\boldsymbol{x}, e_{ki})],$$

where the notation $\Delta p_k(\boldsymbol{x}, e_{ki})$ should be intuitive. Differences in the bottom right term of Equation (21) are therefore how the individual correlation between idiosyncratic preference levels e_{ki} and *changes in* sorting ΔI_{ki} (or $\Delta p_k(\boldsymbol{x}, e_{ki})$) differ by worker type $\boldsymbol{x}$. It seems impossible to sign this term without fully specified parameters and probability distributions in the model. However, ex ante there is no reason why this should be large compared to the other, economically clearly interpretable, terms of Equation (21). Empirically, the $\sum_k \boldsymbol{b}_k'\boldsymbol{x} \cdot \Delta p_k(\boldsymbol{x})$ control in the Propensity Regression partially accounts for $\sum_k \mathrm{E}[e_{ki}\Delta I_{ki} \mid \boldsymbol{x}]$ because it also interacts $\Delta p_k(\boldsymbol{x})$ with $\boldsymbol{x}$ (also see Section E.3).

In the Appendix, I report on Monte Carlo experiments implementing the propensity estimation for Equation (21) (i.e., extending Regression (20) with the interaction of $\Delta p_k(\boldsymbol{x})$ and $\boldsymbol{x}$). The results indicate that, controlling for $\sum_k \boldsymbol{b}_k'\boldsymbol{x} \cdot \Delta p_k(\boldsymbol{x})$, the task prices can be estimated consistently and precisely even when the amenities are talent-specific and workers have idiosyncratic preferences for tasks. However, the amenity coefficients are very imprecise in this case, which is due to the high multicollinearity between the $\Delta p_k(\boldsymbol{x})$ and $\boldsymbol{x} \cdot \Delta p_k(\boldsymbol{x})$ controls as well as the idiosyncratic preferences for tasks.[19] Nonetheless, the main focus of the Propensity Method is to correctly estimate the changing task prices, and this succeeds very well, even in the most general version of the Roy model with idiosyncratic amenities.

3.5 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 existing reduced-form as well as

[18] The choice indicator in period t is $I_k(\boldsymbol{U_{it}}) = \mathbf{1}[\max_{l=1,\ldots,K}\{U_{lit}\} = U_{kit} = \pi_{kt} + \boldsymbol{\beta}_k'\boldsymbol{x_i} + u_{ki} + a_{kt} + \boldsymbol{b}_k'\boldsymbol{x_i} + e_{ki}]$ and the propensity becomes $p_k(\boldsymbol{x}, \boldsymbol{\pi_t}, \boldsymbol{a_t}, \boldsymbol{b}) \equiv \mathrm{E}[I_k(\boldsymbol{U_{it}}) \mid \boldsymbol{x_i} = \boldsymbol{x}]$.

[19] It is know from regression anatomy (i.e., the Frisch-Waugh-Lovell theorem) that this multicollinearity need not affect the estimates ($\Delta\pi_k$) on the other regressors ($\bar{p}_k(\boldsymbol{x})$) in general, though.

structural methods in that it exploits a new 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) and Cavaglia and Etheridge (2017) use panel data in order to control for workers' task-specific fixed effects and observable characteristics (especially experience) in a wage regression. They then assume that conditional task switching is unrelated to changes in the remaining idiosyncratic wage components across tasks and identify the change in prices via time-varying task intercepts. This (conditional) "exogeneity" assumption implies that there is no idiosyncratic skill accumulation or learning about abilities that affects potential wages and makes individuals switch tasks (or stay in their task).[20] Another important aspect in panel data approaches generally is the correct modeling and estimation of differential wage growth across tasks over time (i.e., due to skill accumulation) that would have occurred even in the absence of changing task prices. This parallels the assumption in the repeated cross sections of this paper that, given the talents, the 27 year old workers in two different cohorts are comparable in terms of their skills (at least conditional on observable controls such as educational attainment).

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 similar to interpreting the returns coefficients from a

[20]The key part of the assumption is that workers who stay in a given task, and from whose wage growth the task prices are identified, are not systematically (very) different in terms of their idiosyncratic innovations from all workers who started in that task. Empirically, this may often be a quite innocuous restriction.

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.[21]

The strategy of structural methods for estimating task prices is to model the population distribution of unobservable skills (shocks) explicitly. Examples for this are Heckman and Sedlacek (1985), Gould (2002), or Lee and Wolpin (2006). These papers impose a specific distribution of workers' unobserved skills (e.g., joint normality) and then set up a likelihood function or simulate moment conditions 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.[22] While it does require comparability, the Propensity Method proposed in this paper is derived without explicitly specifying a global distribution of workers' skills under a local approximation of their choice indicators.

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). The fact 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 plausible identification-at-infinity argument. This paper's comparability assumption 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.

[21] 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.

[22] On the other hand, the structural models are able to incorporate many other potentially important features. For example, Lee and Wolpin (2006)'s is a dynamic generalized Roy model which allows for, among other things, human capital accumulation and costs of switching tasks.

An alternative method originally due to Heckman, Lochner, and Taber (1998) uses flat spots in workers' life-cycle skill accumulation to estimate changing prices for skill types (e.g., high-school or college graduates) over time. Bowlus and Robinson (2012), for example, invoke economic theory regarding the optimal investment in human capital over the career as well as evidence on slopes in cross-sectional life cycle profiles to identify the flat spot and then apply it to synthetic panel data constructed from the U.S. Current Population Survey. This method could be modified to estimate task prices (i.e., for different occupation groups) using task stayers' wage growth at the flat spot over time.[23]

Yamaguchi (2016) presents another approach that may 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 as relationships between skills and tasks. The identification relies on unobserved skills being expressible 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.[24]

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

[23] In the data used below, this approach would not work, however, since individuals in the NLSY97 are still too young to have reached their flat spot(s).

[24] 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.

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' skill distributions (the most general one in the paper is a combination model of hierarchical together with independent worker-specific skills across occupations).[25]

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 feature of the method proposed in this paper is that it is easily applied to multiple tasks; with the caveat that the talents $\boldsymbol{x}$ need to be sufficiently detailed to fulfil the first-stage assumption. The method is also transparent in terms of which moments in the data it uses for identification (see Equation (16)).

4 Economic Predictions and Data

4.1 Routine-Biased Technical Change as an Application

The Propensity Method's empirical application estimates the changing prices of abstact, routine, and manual tasks over a period of two decades. In particular, the literature on routine-biased technical change (RBTC) 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),

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

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 some additional restrictions in each of the papers, these models have in common that a rising availability of computer capital (RBTC) leads to an increase in abstract and manual non-routine task prices compared to routine task prices:[26]

$$\Delta(\pi_A - \pi_R) > 0 \text{ and } \Delta(\pi_M - \pi_R) > 0, \tag{22}$$

with $\Delta\pi_k \equiv \pi_{k1} - \pi_{k0}$ and now using the index $k \in \{A, R, M\}$. In the following I estimate the changing task prices and examine the RBTC prediction (22) in U.S. data. In addition, Appendix B derives and tests its implications on the relative wage growth of workers over time who start out in the three different tasks at $t = 0$.

Consistent with Sections 2 and 3, I assume that RBTC affects wages only via market prices for tasks ($\Delta\pi_A$, $\Delta\pi_R$, $\Delta\pi_M$), 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 provide to the market.

4.2 Occupational Grouping and Alternative Hypotheses

The 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. 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,

[26] Earlier versions of this paper explicitly derived (22) from the Autor, Katz, and Kearney (2006) model with a general distribution of workers' skills.

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, 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). That the manual group is mainly made up of services occupations is consistent with a relatively strong sorting of verbal talent into it (see below).

Acemoglu and Autor (2011)'s grouping into broad occupations carries limitations. First, the fine variation in tasks across more detailed occupations is aggregated away, 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, 2018).[27] This enables a consistent comparison of the paper's findings with

[27] 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 cate- gory (sales, office/admin, production, and operator/labor occupations); and a manual ca-

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 most 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 indices (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

tegory (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.

for home-based production), this may raise demand for low-skill services occupations that are intensive in non-routine manual tasks. Manning (2004) and Mazzolari and Ragusa (2013) present evidence that is consistent with this channel.[28] However, SBTC that works independent from RBTC would also imply a rising college premium (Acemoglu and Autor, 2011), and Section B.3 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. I also show 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.[29]

In what follows, I refer to the three occupation groups by *abstract*, *routine*, and *manual tasks*. For intuitive reference they are abbreviated by A, R, and M so that the index k takes the values $k \in \{A, R, M\}$ in the empirical application of the Propensity Method.

4.3 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

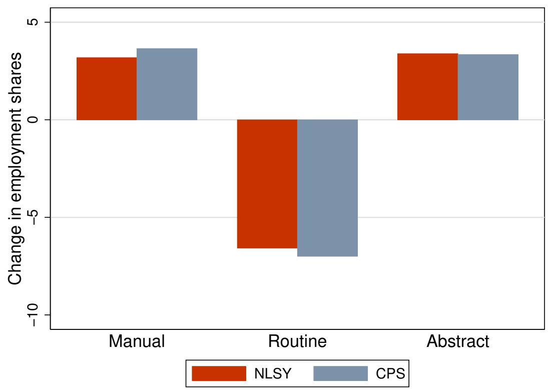
[28] Bárány and Siegel (2018) propose a related explanation in which structural transformation across sectors drives the rising demand for services occupations. Acemoglu and Autor (2011) and Goos, Manning, and Salomons (2014) show that jobs also strongly polarized within industry sectors.

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

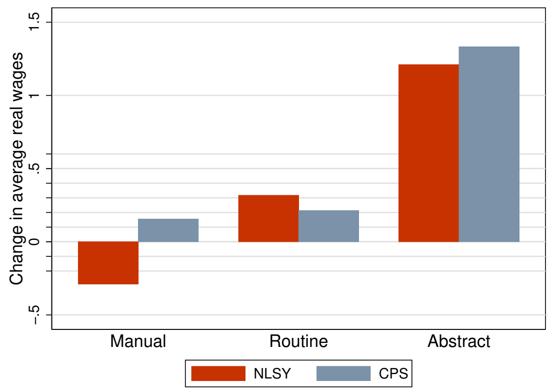
cohorts of the NLSY while male performance remained constant. Therefore, the comparability assumption is more likely to be violated with the available characteristics $\boldsymbol{x}_i$ for females than for males. Moreover, female wages rose substantially across-the-board compared to males and some argue that discrimination against them in different high-skill occupations has declined quite drastically (e.g., Hsieh, Hurst, Jones, and Klenow, forthcoming). Therefore, it is likely that a large part of the returns to characteristics $\boldsymbol{x}_i$ 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 is restricted to males. An earlier version of this paper (Böhm, 2017) summarizes estimates for females.

Figure 1: The distributions of employment and wages for males age 27 in the NLSY and the CPS (1984/92 to 2007/09)

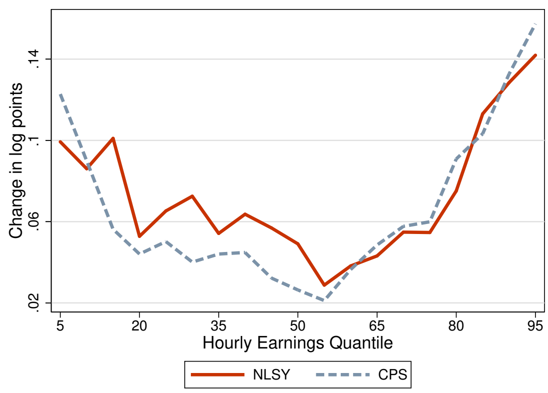
(a) Job polarization



(b) No wage polarization in tasks



(c) Overall wage polarization



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 comparison at constant age arguably reduces the concerns about age effects in earnings (Ashworth, Hotz, Maurel, and Ransom, 2017, further show comparable early work experience in the two NLSYs). The details of the sample construction can be found in Section F of the Appendix.[30]

Figure 1 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. This could be due to changing selection bias into the manual task even if Prediction (22) is true (e.g., see numerical simulations in the Appendix). Finally, the overall wage distribution polarized substantially, both in the NLSY and in the CPS (Panel (c)).

4.4 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 $\boldsymbol{x}_i$ vector, the NLSY talents therefore come as close as possible to fulfilling the comparability Assumption 2.

[30] The sample selection and attrition weighting for the NLSY data is done closely in line with 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 F.1 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.

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 Prada and Urzua (2017) and Speer (2017), who investigate the education and labor market effects of different worker abilities in the NLSY.[31]

Table A.1 in the Online 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 A.1 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).[32] In addition, Table 1 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 $\boldsymbol{x}_i$ vector.[33]

[31]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.

[32]One early determined characteristic that is not constant is the share of Hispanics, which rose by 8 percentage points. Race dummies are part of the vector used for predicting the choice propensities $\bar{p}_k(\boldsymbol{x}_i)$, but I also include the flexibly interacted $\boldsymbol{x}_i$s directly into the respective second specifications of Tables 2, 3, A.3. This controls, among others, for the changing race composition of 27 year old males over time.

[33]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. At the same time, 27 year olds in 2007/09 are recent enough as a cohort to have experienced potentially a substantial

Table 1: Pairwise Correlations between Talents, NLSY 1979 and 1997

	NLSY79 AFQT	Math	Verbal	NLSY97 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		

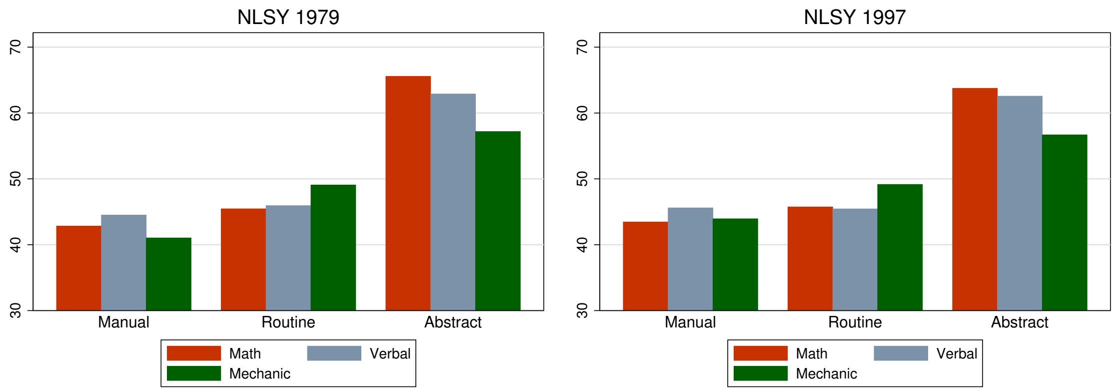
Notes: The table shows the pairwise correlations between composite test scores after standardizing to normal curve equivalents with mean 50 and standard deviation 21.06.

Figure 2 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 governed by their comparative advantage and thus depend on their relative skills. This principle seems to be borne out in Figure 2. 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, which is consistent with it being mainly made up of services occupations (see Section 4.2).

Table A.2 in the Online 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

impact of RBTC (compare Figure 1).

Figure 2: Average talents in tasks, NLSY 1979 and 1997



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 2). 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 A.2 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 Empirical Results

5.1 Task Price Estimates in the Pure Roy Model

Employing the Propensity Regression (18) to the NLSY data estimates the task prices in the pure Roy model:

$$w_{it} = \theta_0 + (\theta_1 - \theta_0)\cdot\overline{p}_A(\boldsymbol{x_i}) + (\theta_2 - \theta_0)\cdot\overline{p}_M(\boldsymbol{x_i}) + \Delta\pi_R\cdot\mathbf{1}[t=1] + \Delta(\pi_A - \pi_R)\cdot\overline{p}_A(\boldsymbol{x_i})\cdot\mathbf{1}[t=1] + \Delta(\pi_M - \pi_R)\cdot\overline{p}_M(\boldsymbol{x_i})\cdot\mathbf{1}[t=1] + \varepsilon_{it}, \tag{23}$$

where I have replaced the general $k \in \{1, ..., K\}$ with the $k \in \{A, R, M\}$ tasks of the empirical application and used the fact that the probabilities sum to one ($\overline{p}_A(\boldsymbol{x_i}) + \overline{p}_R(\boldsymbol{x_i}) + \overline{p}_M(\boldsymbol{x_i}) = 1$) to write (23) as a regression with intercepts in each period. The choice propensities $p_A(\boldsymbol{x_i}, \boldsymbol{\pi_t})$ and $p_M(\boldsymbol{x_i}, \boldsymbol{\pi_t})$ are estimated in first-stage multinomial logit regressions for the NLSY79 ($t = 0$) and NLSY97 ($t = 1$), as discussed in the previous section, and then combined into $\bar{p}_A(x_i)$ and $\bar{p}_M(x_i)$. Thus, for every 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, with the latter requiring the comparability assumption. I also bootstrap the first-stage multinomial choice regressions and the second stage Propensity Regression in order to obtain the correct standard errors given that $\overline{p}_A(x_i)$ and $\overline{p}_M(x_i)$ are estimates with sampling variation.

Table 2 reports the results, showing that the equilibrium prices paid for tasks have changed substantially between the two NLSYs. In particular, in the first row the relative prices of 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. Task prices therefore polarized between the two cohorts of the NLSY. This is consistent with the routine-biased technical change hypothesis, albeit with the qualification that the estimate for the relative manual (and absolute routine) task price is insignificant.[34]

[34]The absolute routine task price is somewhat harder to interpret in the context of routine-biased technical change because RBTC theory makes no clear predictions about absolute task prices (which may, e.g., change due to general productivity growth in the economy too).

The second row alternatively estimates a rich specification for the baseline $t = 0$ as mentioned in Section 3.2. In particular, I add all the regressors used in the first-stage and completely flexibly interact them (i.e., the math, verbal, and mechanical talent tercile as well as race dummies are fully stratified and then interacted with the continuous illicit activities $\times$ precocious sex). The standard errors of the $t = 0$ coefficients strongly rise with these highly correlated regressors and the θ_ks on the baseline $\overline{p}_k(\boldsymbol{x_i})$s become insignificant while the overall R^2 increases (all unreported). However, the coefficients of interest on the interaction between $\overline{p}_k(\boldsymbol{x_i})$ and $\mathbf{1}[t = 1]$ as well as their standard errors hardly change at all. That is, as predicted above, the estimated task price changes in Table 2 are robust to a rich addition of $\boldsymbol{x_i}$.[35]

The remainder of the Table returns to the more parsimonious main specification (23) and examines the robustness of its results when potentially confounding forces impact the wage distribution, motivated by the findings of Section E's Monte Carlo experiments. An important competitor hypothesis to RBTC, discussed in Section 4.2, is a combination of skill-biased technical change together with rising consumption demand for services. Since SBTC implies a rising return to college (Acemoglu and Autor, 2011), I next estimate task prices allowing for independently changing returns to education. In rows three and four of Table 2, an augmented Propensity Regression is estimated, which accounts for shifting education composition as well as educational premia that change independently of RBTC, controlling for college and degree dummies interacted with time. The task price changes remain qualitatively the same in these estimations, with the relative price of the manual task somewhat increasing and the (unreported) rise of the college premium is insignificant. The price polarization result thus persists when allowing for changing composition or rising returns to education (for further results in favor of RBTC as opposed to SBTC see Section B.3).

In row five of Table 2 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 3.2 and it is explained in detail in the Appendix. While less straightforward to implement, an advantage of the OMD is that it provides an overidentifying restrictions test ("J-test") of the moment

[35]Unreported Monte Carlo simulations confirm this result.

Table 2: Estimated Task Price Changes in the NLSY (1984/92 to 2007/09)

	$\Delta(\pi_A - \pi_R)$ in log points (s.e.)	$\Delta(\pi_M - \pi_R)$ in log points (s.e.)	$\Delta\pi_R$ in log points (s.e.)	Model Test (p-value in %)
OLS on Propensities	25.1 (12.2)	32.9 (39.6)	-4.2 (7.4)	
OLS on Propensities (t=0: all $\mathbf{x_i}$ interacted)	25.9 (13.4)	33.8 (38.8)	-4.8 (7.6)	
OLS on Propensities (2nd-stage college)	19.5 (14.2)	46.6 (39.9)	-6.4 (7.6)	
OLS on Propensities (2nd-st degree dum.)	25.6 (14.6)	56.9 (40.2)	-5.2 (9.5)	
Opt. Min. Distance	20.2 (6.6)	38.9 (26.4)	-3.2 (3.3)	12.3 (13.8)
OLS on Propensities (Adj. for min. wage)	27.3 (12.6)	32.0 (41.0)	-5.7 (7.7)	

Notes: The first row of the table presents estimated task price changes from the basic (23). Row two adds x_i to the Propensity Regression baseline and fully interacts its elements with each other. The third and fourth row run the augmented Propensity Regression (18′), adding college and degree dummies interacted with time respectively. Row five reports the task price changes from a minimum distance estimation explained detail in the Appendix. 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.

conditions implied by the empirical model, which is reported in the last column of row four. The (asymptotic) standard errors are also lower due to the OMD's optimality. Reassuringly, the J-test does not reject the model and the point estimates in row four of Table 2 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 1 and, since the manual task workers are more frequently found in this lower end, it may distort the task price estimates. The Monte Carlos Section E.2

shows 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.[36] The wage distribution for 27 year olds in the NLSY, and for comparison the CPS, is now substantially flatter in the bottom than without the adjustment (compare the solid lines in Panels c and d of Figure 3 below). Row six of Table 2 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.

In Appendix B, I further examine the wage growth over time as a function of different propensities $p_k(\boldsymbol{x_i}, \boldsymbol{\pi_0})$ to start out in the three tasks at $t = 0$. Consistent with Prediction 3 (also derived in the Appendix), wage growth of abstract as well as manual starters is substantially larger than that of routine starters, which is actually negative though insignificant. Again, this result is robust to specifications allowing for changing

[36] 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 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

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

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.

returns to other skills.

5.2 Task Price Estimates in the Generalized Roy Model

The following reports estimation results for the case with homogeneous amenities of Section 3.3. Analogous to the pure Roy case, I rewrite Equation (20) as the actual regression that is estimated:

$$\begin{aligned} w_{it} = \theta_0 &+ (\theta_1 - \theta_0)\overline{p}_A(\boldsymbol{x_i}) + (\theta_2 - \theta_0)\overline{p}_M(\boldsymbol{x_i}) + (\psi_1 - \psi_0)\Delta p_A(\boldsymbol{x_i}) + (\psi_2 - \psi_0)\Delta p_M(\boldsymbol{x_i}) \\ &+ \Delta\pi_R \cdot \mathbf{1}[t=1] + \Delta(\pi_A - \pi_R)\cdot\overline{p}_A(\boldsymbol{x_i})\cdot\mathbf{1}[t=1] + \Delta(\pi_M - \pi_R)\cdot\overline{p}_M(\boldsymbol{x_i})\cdot\mathbf{1}[t=1] \\ &+ (-\bar{a}_A + \bar{a}_R)\cdot\Delta p_A(\boldsymbol{x_i})\cdot\mathbf{1}[t=1] + (-\bar{a}_M + \bar{a}_R)\cdot\Delta p_M(\boldsymbol{x_i})\cdot\mathbf{1}[t=1] + \varepsilon_{it}, \end{aligned} \tag{24}$$

where Δp_A and Δp_M are the difference over time of the propensities estimated in the first-stage and I have made use of the fact that $\Delta p_R(\boldsymbol{x_i}) = -[\Delta p_A(\boldsymbol{x_i}) + \Delta p_M(\boldsymbol{x_i})]$. The rest of the implementation is the same as for the pure Roy case, although one thing to notice once again is that only the relative amenities can be identified.

Table 3 reports the results from this estimation. The first column shows that the estimates for the relative abstract task price change are slightly smaller and standard errors slightly larger than in the pure Roy, but in sum do not change much at all and are robust to the more general estimation specification. The estimate for the relative price change of the manual task in the second column is larger than before and it becomes statistically significant at around the ten percent level in rows three and four. The point estimates of the price changes for the routine task (third column) remain small and statistically indistinguishable from zero in the generalized estimates. Table 3's alternative specifications show largely similar coefficients to the first row and to those in the pure Roy estimation of Table 2. A partial exception are the optimal minimum distance results, which are substantially lower but nonetheless indicate (in terms of point estimates) positive relative price changes for the abstract and manual tasks. Therefore, these results imply that the task price estimates are largely robust, in terms of point estimates and statistical significance, to the extension of the estimation method that accounts for homogeneous amenities as in Section 3.3.

Table 3 also reports workers' average non-pecuniary valuations of different tasks. The second to last column presents the valuations for the abstract relative to the routine task. The point estimates for these are about 50–55 log points, which is substantial (e.g., the increase in the abstract task price between the two sample periods is about half that number), although only statistically significant in one of the specifications. The valuation of the manual compared to the routine task is even larger in terms of point estimates, with a very high estimate for the optimal minimum distance specification (row 4).[37] This is statistically significant around the ten or five percent levels in the different specifications, but also with quite wide confidence intervals.[38]

The estimated amenity values can be understood by considering the difference in the results of Appendix B and the pure Roy Table 2 above. In Table B.2, estimated relative wage growth of abstract task starters, using $p_A(x_i, \pi_0)$ as a regressor in Equation (B.3), is higher than the estimated relative abstract task price growth of Table 2. Also, using $p_M(x_i, \pi_0)$ as a regressor in (B.3), wage growth of manual task starters is substantially higher than the estimated manual task price growth of Table 2. The reason for this is that those talent types x who move out of routine and into abstract (and thus have high $p_A(x, \pi_1)$ compared to $p_A(x, \pi_0)$) and especially into manual (high $p_M(x, \pi_1)$ compared to $p_M(x, \pi_0)$) tasks experience lower wage growth. According to the generalized Roy model of Section 2, these moves together with the low *conditional* wage growth can be rationalized by the fact that the movers gain amenities in their new tasks, the implied values of which are reported in the last two columns of Table 3.

An alternative way to interpret the amenity together with the task price results is to consider the effects for different types of workers in terms of x characteristics. According to the pure Roy model, rising abstract prices benefit high-math talented individuals who sort strongly into that task (Table A.2). The pure Roy effect also implies

[37] Sorkin (2018, Figure 5b) finds variation in non-pecuniary valuations of one-digit industries comparable to the Propensity Regressions. The OMD estimate is admittedly quite extreme, however. The reason for this may be the outsize influence of particular moment conditions, as equally weighted minimum distance (i.e., OLS on the moment conditions instead of inverse weighting by their estimated variance, see Appendix) yields a somewhat larger estimate for the relative abstract and manual task price changes as well as a slightly higher abstract and a lower manual amenity (not reported). Moreover, despite its (asymptotic) optimality, the standard errors in the OMD are quite large here, too.

[38] The Monte Carlo Section E.3 indicates that amenities are in general less precisely estimated than task prices in the Propensity Regression.

Table 3: Task Price Changes and Average Amenities in the NLSY (1984/92 to 2007/09)

	$\Delta(\pi_A - \pi_R)$ log points (s.e.)	$\Delta(\pi_M - \pi_R)$ log points (s.e.)	$\Delta\pi_R$ log pts (s.e.)	$\bar{a}_A - \bar{a}_R$ amenity (s.e.)	$\bar{a}_M - \bar{a}_R$ amenity (s.e.)
OLS on Propensities	21.9	50.9	-1.4	41.7	78.2
	(13.5)	(45.2)	(8.0)	(36.0)	(43.2)
OLS on Propensities	25.7	44.7	-1.6	69.5	59.3
(t=0: all $\mathbf{x_i}$ interactd)	(15.4)	(44.7)	(8.8)	(43.6)	(44.1)
OLS on Propensities	17.4	67.2	-3.6	51.1	81.9
(2nd-stage college)	(14.9)	(41.3)	(7.9)	(35.2)	(42.0)
OLS on Propensities	24.3	77.2	-4.4	61.5	68.2
(2nd-st degree dum.)	(15.3)	(41.8)	(9.7)	(34.6)	(41.9)
Opt. Min. Distance	11.3	26.8	4.7	14.6	138.5
	(10.5)	(32.6)	(5.6)	(50.2)	(63.8)
OLS on Propensities	23.7	51.1	-2.7	44.7	84.6
(Adj. for min. wage)	(13.9)	(47.0)	(8.4)	(36.9)	(44.0)

Notes: The table shows the results from the estimation method for the generalized Roy model with homogeneous amenities. The first row presents task price changes and amenity values from the Propensity Regression (24). Row two adds x_i baseline and fully interacts its elements with each other. The third and fourth row add college and detailed degree dummies interacted with time respectively. Row five reports the task price changes from the optimal minimum distance estimation explained in the Appendix (model test p-value: 71.8). 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 s.e. (500 iterations) below the coefficients.

that rising abstract and manual prices would raise wages of high-verbal workers who sort into both of these tasks. However, at the same time the sorting of verbal talent into abstract as well as manual tasks further rises over time (again, see Table A.2). This comes with increasing amenities enjoyed by high-verbal workers from working in those tasks, and thus with a decline in their wages predicted by the generalized Roy model (but of course a further rise of utility, which cannot be directly measured). Indeed, unreported reduced-form wage regressions show that the return to math talent rises substantially between the NLSY79 and NLSY97 whereas the changing returns to verbal are not very different from the returns to mechanical talent. The generalized Roy model therefore rationalizes changing talent returns over time using both the effects implied by task prices (*sorting*) and by the amenities (*re-sorting*).

The high amenities of the manual task in Table 3 also provide a new explanation for the generally low wages in that task. In particular, the literature on job polarization has somewhat struggled to explain why workers have moved from higher-earning routine to lower-earning manual tasks, which include many low-skill services occupations, over time.[39] According to the analysis at hand, the higher wages that individuals are observed to earn in the routine relative to the manual task are to a substantial degree due to compensation for disamenities. At the same time, observed wage differences between abstract and routine would be even larger were it not for the higher amenities in the abstract task. Therefore, the job polarization trends over time are indeed consistent with demand shocks for tasks, as relative prices and employment co-move, and at the same time with switches from routine to manual tasks being associated with falling wages due to rising non-pecuniary amenities.

Finally, I estimate the most general model with heterogeneity in non-pecuniary task valuations of Section 3.4, which adds controls $\boldsymbol{x_i}\cdot\Delta p_A(\boldsymbol{x_i})$ and $\boldsymbol{x_i}\cdot\Delta p_M(\boldsymbol{x_i})$ interacted with time to Propensity Regression (24). Table A.3 in the Online Appendix reports the results. The point estimates for the change of the relative abstract task price remain again remarkably stable at around 20–25 log points in the different specifications (the exception being row two with fully interacted $\boldsymbol{x_i}$ baseline), although they are not statistically significant anymore. However, the relative manual task price changes now drop to about -8 to +28 log points and are far away from statistical significance. This indicates that, in the NLSY data, the $\boldsymbol{x}$ variables are not powerful enough in the case of the manual task to sufficiently distinguish between $\bar{p}_k(\boldsymbol{x})$ and $\boldsymbol{b}_k'\boldsymbol{x}\cdot\Delta p_k(\boldsymbol{x})$ for the price estimates to be unaffected by the generalization to heterogeneous non-pecuniary valuations across $\boldsymbol{x}$-types.[40] The estimates for the amenity intercepts (and the unreported talent loadings $\boldsymbol{b_k}-\boldsymbol{b_R}$ on tasks) also become very imprecise, and in the case of the manual amenity $\bar{a}_M-\bar{a}_R$ ($\bar{a}_A-\bar{a}_R$ for the fully interacted $\boldsymbol{x_i}$) quite extreme in terms of point estimates.[41]

[39]For example, Autor and Dorn (2013) allow for the possibility of falling wages in the low-skill services occupations by letting the overall effect on wages depend on the substitutability between goods and services in final consumption (and the substitutability of routine labor and capital) compared to the importance of the routine component in goods production. Also, Cortes (2016) finds that workers who switch from routine to manual tasks in panel data initially experience a significant wage drop.

[40]The Appendix Monte Carlo simulations however indicate that, with the right data, the correct task price changes can in principle be identified even in the heterogeneous amenity case.

[41]Appendix D shows that it is very difficult to implement the optimal minimum distance estimation

These results suggest that on the one hand the heterogeneous non-pecuniary specification based on (21) may be too rich for what is estimable in the (limited) NLSY data as the talents $\boldsymbol{x}$ are used in multiple regressors, which introduces quite a bit of multicollinearity. On the other hand, the coefficients $\bar{a}_k - \bar{a}_R$ do not have such a central interpretation in the heterogeneous amenity case anymore (they are only intercepts of the non-pecuniary utility Equation (3)), and they cannot be consistently estimated in the flexible generalized Roy model anyway (see results and discussion in Monte Carlo Section E.3). In the end, the predominant interest of the Propensity Method is to identify the correct task prices and there the results are mixed: while the relative abstract price changes are largely stable in this most flexible specification, the relative manual estimates do change.

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:

$$\widehat{w_{i1}^{TP}} = w_{i0} + \widehat{\Delta\pi_R} + I_A(\boldsymbol{U_{i0}})\Delta(\widehat{\pi_A - \pi_R}) + I_M(\boldsymbol{U_{i0}})\Delta(\widehat{\pi_M - \pi_R}) \tag{25}$$

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 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

in the heterogeneous amenities case. The OMD is therefore left out from Table (A.3).

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.

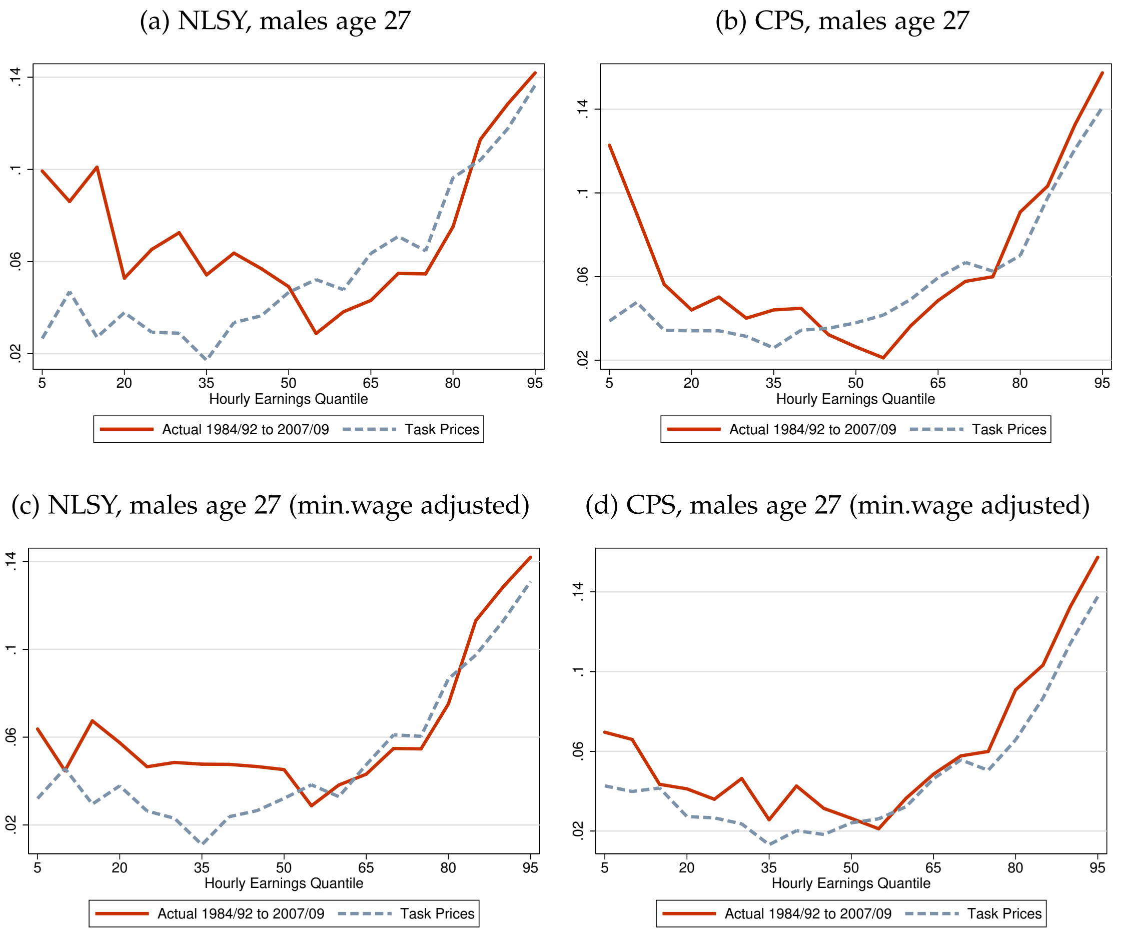
The top row of Figure 3 plots the predicted change in the wage distribution (the quantiles of the distribution of $\widehat{w_{i1}^{TP}}$ minus the respective quantiles of the distribution of w_{i0}) together with the actual change in the wage distribution ("w_{i1} minus w_{i0}") for 27 year olds in the NLSY and CPS. The task price changes used in the predicted wage distribution, also in the CPS, are from the first specification in the pure Roy estimation of Table 3 (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. Figure A.2 in the Online Appendix shows the same plots using the task prices from the generalized Roy estimation in row one of Table 2.[42] The results are qualitatively similar, which indicates that at least the generalization to homogeneous amenities does not alter the conclusion about the potential effect of task prices on the 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, Figure A.1 in the Online Appendix 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

[42]The average predicted wage change in Figure A.2 is adjusted downward to match the average actual wage change. This is necessary because, in the eyes of the model, the fact that workers move into higher-amenity abstract and manual tasks over time reduces their wages. Therefore, the estimated task prices have to be higher in levels in order to match observed wage growth. In that sense, one can compare the shapes of the predicted with the actual wage distribution in Figure A.2, but not their levels.

Also, the predicted change of the wage distribution for the heterogeneous amenities model is left out, since the manual task price estimates seem not robust to this most general estimation.

Figure 3: 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)



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 is also discussed in relation to theoretical Prediction 2 in the Appendix: 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 multi-dimensional skill model. Figure A.3 in the Online Appendix 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 of wages in the upper half and weakens the increase of wages in the lower half of the distribution when task prices polarize.[43]

As discussed in Section B.1, 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 Figure A.3 in the Online Appendix 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 B.1). The polarizing task prices and the changing wages of workers who start out in different tasks, as estimated in this paper, 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 in the lower half 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 3 plots the actual and the predicted distribution when wages are adjusted for the change in the real value of the minimum wage and the task price estimates are taken from the bottom specification in Table 2. The fit in the lower half 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. Again, the results for the generalized Roy estimation in Figure A.2 in the Online Appendix are similar.

[43] 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 B.1 Panel (d)).

6 Conclusion

This paper proposed a new method to estimate task price changes per efficiency unit of skill as well as non-pecuniary amenities across tasks. I showed theoretically that, in the generalized Roy (1951) model, workers' wage growth over time exclusively depends on the relationship of their task choices with amenities and these changing prices; and not on skill levels or any particular distribution of skills in the economy. I then devised an approach to empirically implement this insight by leveraging the fact that observable talents will experience differential returns over time depending on the changes in prices (and the amenities) of those tasks that they predict workers to sort into.

Applying this new Propensity Method to data from the U.S. National Longitudinal Survey of Youth shows that, consistent with routine-biased technical change (RBTC), task prices polarized 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. The estimation also reveals that non-pecuniary amenities enjoyed in abstract and manual tasks are substantially higher than in routine tasks. Finally, the findings suggest that task price changes have led to a widening of inequality in the upper half of U.S. males' wage distribution and to a flattening in the lower half, but appear unable to explain all of the increase that is observed in the bottom.

This paper's new theoretical result in the generalized Roy model and the Propensity Method to estimate task prices could be 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.

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.

Ashworth, J., V. J. Hotz, A. Maurel, and T. Ransom (2017): "Changes across cohorts in wage returns to schooling and early work experiences," Discussion paper, National Bureau of Economic Research.

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.

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. L., and C. Siegel (2018): "Job polarization and structural change," *American Economic Journal: Macroeconomics*, 10(1), 57–89.

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.

Böhm, M. J. (2017): "The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change," *IZA Discussion Paper 11220*.

Bowlus, A. J., and C. Robinson (2012): "Human Capital Prices, Productivity, and Growth," *American Economic Review*, 102(7), 3483–3515.

Cavaglia, C., and B. Etheridge (2017): "Job Polarization, Task Prices and the Distribution of Task Returns," Working Paper.

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.

Cortes, G. M., and G. Gallipoli (2017): "The costs of occupational mobility: An aggregate analysis," *Journal of the European Economic Association*, 16(2), 275–315.

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.

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 (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.

Gould, E. D. (2002): "Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States," *Journal of Labor Economics*, 20(1), 105–147.

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., L. Lochner, and C. Taber (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58.

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 (forthcoming): "The Allocation of Talent and U.S. Economic Growth," *Econometrica*.

Keane, M., and K. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

Lee, D., and K. I. Wolpin (2006): "Intersectoral labor mobility and the growth of the service sector," *Econometrica*, 74(1), 1–46.

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*.

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. Urzua (2017): "One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance, and Earnings," *Journal of Labor Economics*, 35(4), 953–991.

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

Sorkin, I. (2018): "Ranking firms using revealed preference," *The Quarterly Journal of Economics*, 133(3), 1331–1393.

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.