Routine-Biased Technical Change: Panel Evidence of Task Orientation and Wage Effects*

Matthew B. Ross†

Ohio State University

Keywords: Labor Dynamics, Occupations, Routine-Biased Technical Change, Technological Change, Tasks, Wage and Employment Polarization

JEL: J23, J24, J31, J62, J63, J64, O31, O33

Abstract: This analysis explores the impact of Routine-Biased Technical Change (RBTC) on employment and wages using a combined panel of occupational tasks and individual workers. The existing literature on RBTC has relied exclusively on cross-sectional measures of occupational task content and has been unable to directly examine changes in relative task premiums or the consequent sorting of workers. Using incumbent-updated survey data from a series of archived releases of the Occupational Information Network (*O*NET*) database, I develop a novel methodology for constructing a panel of occupational tasks. Exploiting variation in task content within occupations over time, I find new and compelling evidence in support of the RBTC hypothesis. More specifically, I find that the relative premium for routine tasks has declined more rapidly in routine-intensive occupations and that self-selection has increased sharply over time.

* I am thankful for insightful comments from Kenneth Couch, Delia Furtado, R. Kaj Gittings, Jesse Kalinowski, Subhash Ray, Stephen L. Ross, and Austin Smith. This paper also benefitted from constructive discussion by participates at the Federal Reserve Bank of Atlanta seminar series, IZA/CEDEFOP Workshop on Skills and Skills Mismatch, 2015 Southern Economic Association meetings, 2016 Western Economic Association meetings, and the University of Connecticut seminar series.

† Corresponding Author: M.B. Ross, Department of Economics; Ohio State University; 1945 North High Street; Columbus, OH 43210. Ross.1469@OSU.edu

This analysis fills a significant gap in the empirical literature on Routine-Biased Technical Change (RBTC) by directly examining wage and employment dynamics using a combined panel of occupational tasks and individual workers. In examining wage and employment polarization, the existing literature on RBTC has relied exclusively on cross-sectional measures of occupational task content. In this paper, I exploit occupational variation in task content over time for identification and develop a natural extension to Autor and Handel (2013) using panel data. Panel data on occupational task content allows for further testing of the RBTC hypothesis by examining the consequent sorting of workers based on comparative advantage as well as changes to relative task premiums over time. Further, I am able to fully controls for time invariant factors like unobserved heterogeneity through estimation with individual, occupation, and job-spell fixed-effects. In strong support of the narrative of wage and employment polarization that is posited by RBTC, I find that the relative premium for routine tasks has declined more rapidly in routine-intensive occupations and that self-selection has increased sharply over time.

This paper proceeds along the following outline: The first section provides an introduction and overview of the relevant literature. The second section details an extension of the existing theory underlying the RBTC hypothesis and derives several empirically testable implications. The third section details the construction of a synthetic panel of occupational task content and provides descriptive statistics from that data as well as the panel of individual workers. The fourth section provides an empirical analysis of wage effects and self-selection. The fifth section makes progress towards causal identification by applying instrumental variables estimation. The final section summarizes the findings and provides some conclusory remarks.

1. Introduction

In describing why technology has resulted in some occupations becoming more automated than others, a recent paper by David Autor (2014) outlines a compelling mechanism for observed changes in the labor market. In this paper, Autor refers to tasks that follow explicit rules as routine and suggests that they are more easily codified by technology. Codification of these tasks allows for them to be more easily substituted for capital in the production process. In contrast, tasks that are rich in tacit knowledge are characterized as non-routine. These tasks are less easily codified because they require frequent use of cognitive judgment or social interaction. Non-routine tasks, unlike routine tasks, utilize capital as a complement in production.

This nuanced view of technological change suggests that the primary driving force behind observed changes in the labor market is the falling price of computing power coupled with the increased capability of technology to replicate human tasks. More specifically, these factors have displaced workers in occupations with a high degree of content in routine tasks while simultaneously increasing the demand for workers engaged in non-routine tasks. Empirical evidence of this predicted pattern of displacement and wage polarization has been prominently documented in works by Katz and Murphy (1992); Autor, Katz, and Krueger (1998); Autor, Levy, and Murnane (2003); Autor, Katz, Kearney (2005); Acemoglu and Autor (2011).

Acemoglu and Autor (2011) develop a theoretical model of RBTC that provides a comprehensive exposition of the interconnectedness between technology, tasks, skills, and wages. A key feature of their model is the distinction made between employers' demand for tasks and workers' supply of skills. The model structures production as a function of routine and non-routine task where occupations are distinct bundles of these labor inputs. Skills, on the other hand, determine the efficiency of a worker at completing a given task and are exogenously determined by ability or attainment of human capital. Thus, the labor market is characterized by an imperfect matching of skills to tasks and the requisite sorting of workers across occupations based on comparative advantage. The model uses a fully developed supply and demand framework to derive comparative statics related to task replacing technology, an important characteristic of the RBTC hypothesis. The model has been subsequently expanded to accommodate empirical applications in a stream of literature that has recently been characterized as taking a task-based approach.

Motivated by Acemoglu and Autor (2011), Firpo, Fortin, and Lemieux (2013) develop a cross-sectional Roy model that is used to examine the distribution of wages within occupations. The application of a Roy model accommodates a task-based framework and allows for the cross-occupation transferability of skills described (See Gathmann and Schönberg 2010). Autor and Handel (2013) apply a similar Roy model to cross-sectional survey data on self-reported task engagement. The authors combine occupation-level task measures with self-reported task inputs and use the interaction to account for potential self-selection into occupations. Altonji, Kahn, and Speer (2014) use a similar framework to investigate the forces behind changes to the wage distribution across college graduates from different fields of study. Each of these analyses have document important aspects of wage and employment polarization using cross-sectional data on occupational task content. [1]

Recent work by Cortes et al. (2014; 2016) links cross-sectional measure of task content to panel data on individual workers and examines the employment and wage dynamics of those initially employed in routine-intensive occupations. Cortes (2016) finds evidence that workers with high ability are more likely to switch into non-routine occupations and that workers with low ability have a higher probability of switching to occupations dominated by non-routine tasks. In examining transition rates using task variation across occupations, Cortes et al. (2014) details empirical evidence that an increase in the transition rate from non-employment to employment coupled with a decrease in the transition from employment to non-employment has played a crucial role in the disappearance of routine jobs. Additional evidence by Michael Böhm supports these findings and reports that RBTC can help explain much of the changes in U.S. wage inequality over the last two decades.

The existing empirical literature on RBTC has been limited to examining of wage and employment polarization using cross-sectional data on occupational task content. Autor and Handel (2013) use a self-reported cross-section of task engagement to test an integral component of the RBTC hypothesis. Specifically, the authors use of self-reported data on task engagement allows them to test for the comparative advantage that driving self-selection across occupations. The panel data on occupational task content assembled in this analysis allows for further testing of the model outlined by Autor and Handel as well as additional components of the RBTC hypothesis. In particular, panel data allows for an examination of whether the task premium for routine task engagement has fallen more

[1] Related work includes Blender (2007), Jensen and Kletzer (2010), Yamaguchi (2011), and Cortes et al. (2014; 2016)

sharply in routine-intensive occupations and whether the impact that this has had on consequent sorting of workers based on comparative advantage. In estimating associated wage effects, panel data also has the advantage of allowing a two-way fixed-effects framework that controls for things like unobserved individual ability and time invariant occupational premiums.

There are two notable exceptions where authors rely on German panel data that includes self-reported levels of task engagement, Spitz-Oener (2006) and Gathmann and Schönberg (2010). Although distinct from my analysis in both purpose and scope, these analyses provide additional support for using within occupation variation in task content for identification. In particular, Spitz-Oener (2006) examines changes in reported task engagement both within and across occupations over a twenty-year period and relates these changes to technological change. She finds evidence that the most significant changes in task engagement have occurred in occupations that have experienced a rapid adoption of computer technology since 1979. Using the same data, Gathmann and Schönberg (2010) explore the differences between task-specific (semi-portable) occupational skills and more general forms of human capital. The authors find evidence that individuals are more likely to transition to an occupation with similar task engagement to their source occupation and that patterns of wage growth persist through these transitions.

In this paper, I construct a similar dataset to that used by Spitz-Oener (2006) and Gathmann and Schönberg (2010) but focuses more directly on examining the RBTC hypothesis through an examination of wage and employment dynamics. I develop a natural extension to Autor and Handel (2013) that allows for further testing of the wage dynamics and self-selection associated with RBTC. More specifically, I use the combined panel to examine changes in the relative premium paid for routine task engagement and the consequent sorting of workers based on comparative advantage. Since identification comes from within occupation variation in task content over time, the empirical analysis can control for time invariant factors like unobserved heterogeneity through estimation using individual, occupation, and job-spell fixed-effects. I find strong evidence that the relative premium for routine tasks has declined most rapidly in routine-intensive occupations and that self-selection has increased sharply over time. These findings provide compelling new evidence that the RBTC hypothesis accurately describes the mechanisms driving the wage and employment polarization observed in the labor market.

2. Theoretical Model

Following the recent literature on RBTC, I develop a task-based models of the labor market and derive several implications related to wages and self-selection. The motivation and structure of the model is similar to Autor and Handel (2013) but includes additional important elements from Cortes (2016) and Peri and Sparber (2009). Rather than modeling tasks as a continuum, I simplify Autor and Handel's framework by considering only routine and abstract task engagement (see Peri and Sparber 2009). I also assume that occupational production takes a Constant Elasticity of Substitution (CES) form rather than the less general Cobb-Douglas setting (see Peri and Sparber 2009; Cortes 2016). The advantage of this combined framework is that it allows for a direct examination of the implications of RBTC as they relate to wage and employment dynamics.

2.A Theoretical Model: Production and Occupational Choice

The efficiency of worker i in completing abstract and routine tasks is represented by $\phi_{i,a}$ and $\phi_{i,r}$ respectively. Rather than considering a continuum of tasks and skill as in Autor and Handel (2013), a worker's skill endowment is written in two dimensions such that $\Phi_i = \{\phi_{i,a}, \phi_{i,r}\}$ where $\phi_{i,a}$ and $\phi_{i,r}$ are strictly positive numbers with a continuous support. Task efficiencies are normalized for a given time interval and are assumed to be exogenous, that is, the product of innate skill or prior investments in human capital and formalized training. Following Peri and Sparber (2009), I represent the share of time a worker spends performing abstract tasks as $\ell_{t,i}$ and the share performing routine tasks as $(1 - \ell_{t,i})$. A worker i's supply of abstract and routine tasks, in terms of efficiency units, is equal to $a_{t,i} = (\ell_{t,i})^{\delta}\phi_{i,a}$ and $r_{t,i} = (1 - \ell_{t,i})^{\delta}\phi_{i,r}$ respectively. The parameter $\delta \in (0,1)$ allows for decreasing returns from task engagement and prevents the possibility that an individual fully specializes.

Next, I consider an economy where abstract and routine tasks are combined in a CES production structure to produce one unit of output for occupation j in time period t such that:

$$y_{t,j} = \left[\beta_j A_{t,j}{}^{\rho} + (1-\beta_j)(K_t \mathrm{R}_{t,j})^{\rho}\right]^{\frac{1}{\rho}} \qquad 1)$$

In this economy, the final consumption good Y_t is composed of a linear combination of output $Y_t = \sum_j \rho_j y_{t,j}$ for all occupations.[2] The variable $A_{t,j}$ represent the aggregate abstract task input for all workers in occupation j such that $A_{t,j} = \sum_{i \in j} a_{t,i} = \sum_{i \in j} (\ell_{t,j,i})^{\delta} \phi_{i,a}$ and $\mathrm{R}_{t,j}$ represents aggregate routine task input such that $R_{t,j} = \sum_{i \in j} r_{t,i} = \sum_{i \in j} (1-\ell_{t,j,i})^{\delta} \phi_{i,r}$. Following Cortes (2016), the variable K_t represents exogenously given capital stock that acts as a factor productivity shifter for aggregate routine task input. The coefficient β_j is the occupation-specific share of abstract task content required to produce one unit of output. The elasticity of substitution between abstract and routine tasks is simply $\sigma = \frac{1}{(1-\rho)}$ where $\rho \in (-\infty, 1)$ and holds across all occupations.

Assuming a competitive labor market and perfect competition, the occupation-specific task premium for abstract relative to routine task inputs can be written as:

$$\frac{\lambda_{t,j,a}}{\lambda_{t,j,r}} = \frac{\beta_j}{(1-\beta_j)} \left(\frac{A_{t,j}}{K_t \mathrm{R}_{t,j}}\right)^{\rho-1} \qquad 2)$$

A worker i in occupation j, taking the occupation-specific task premiums as given, receives log wages such that:

[2] Although I assume here that the intermediate good is additively separable, nesting occupational production within another CES framework will not have an impact on the implications of the model.

$$w_{t,j,i} = \alpha_j + \lambda_{t,j,a}\left(\left(\ell_{t,j,i}\right)^{\delta}\phi_{i,a}\right) + \lambda_{t,j,r}\left(\left(1-\ell_{t,j,i}\right)^{\delta}\phi_{i,r}\right) \quad 3)$$

A worker i's wage in occupation j in time period t is a function of occupation-specific wage premiums in abstract and routine task engagement. The parameter α_j is not constrained to be positive and represents a premium of penalty that captures output elasticity of the final good, market demands, or a possible skill floor. Assuming that output has been normalized to unity across occupations, the production structure for a single occupation j can be summarized by $\Lambda_{t,j} = \{\alpha_j, \lambda_{t,j,a}, \lambda_{t,j,r}\}$.

A worker, taking the production structure each period as given, selects into an occupation j that maximizes earnings such that:

$$w_{t,i} = \max_j\left\{\alpha_j + \lambda_{t,j,a}\left(\left(\ell^*_{t,j,i}\right)^{\delta}\phi_{i,a}\right) + \lambda_{t,j,r}\left(\left(1-\ell^*_{t,j,i}\right)^{\delta}\phi_{i,r}\right)\right\} \quad 4)$$

The economy is characterized by self-selection of workers into occupations based on comparative advantage.[3] The intuition of the model is that workers first consider the optimal allocation of their time endowment within each occupations and then self-select into the occupation where they receive the highest market wage given their skill endowment. Taking prices and the skill endowments as given, the optimal labor allocation $\ell^*_{t,j,i}$ in each occupation is the value that maximizes wages and solves the following equality $\ell_{t,j,i}/(1-\ell_{t,j,i}) = \left(\lambda_{t,j,a}\phi_{i,a}/\lambda_{t,j,r}\phi_{i,r}\right)^{1-\delta}$.

2.B Theoretical Model: Routine-Biased Technical Change

In the present model, it is intuitively clear that RBTC will have an impact on both wages and self-selection. To fix ideas, I now explore these implications formally by imposing an exogenous growth in capital stock $K_{t+1} > K_t$ that increases the factor productivity of

[3] Assuming a continuum of occupations would imply that the marginal worker in occupation j in period t is indifferent to the next best alternative but there is no need for such a restrictive assumption.

routine task engagement. We first consider the movement and magnitude of changes to task premiums in response to a change in the factor productivity of routine tasks. In the latter portion of the chapter, I also consider the implications that these changes have on the self-selection of workers into routine and abstract-intensive occupations.

Taking the partial derivative of (2) we find that:

$$\frac{\partial\left(\lambda_{t,j,a}/\lambda_{t,j,r}\right)}{\partial K_t} = (1-\rho)\frac{\beta_j}{\left(1-\beta_j\right)}\left(\frac{A_{t,j}}{K_t \mathrm{R}_{t,j}}\right)^{\rho}\frac{1}{K_t}$$

An exogenous increase in capital stock will cause the relative premium for abstract tasks to increase because the partial derivative of (2) is positive. The increase will have a differential impact depending on the share $\frac{\beta_j}{(1-\beta_j)}$ of abstract relative to routine tasks. Consider two occupations j, k such that $\beta_j < \beta_k$ which implies mathematically that the relative premium for abstract tasks increases more $\Delta\frac{\lambda_{t,j,a}}{\lambda_{t,j,r}} > \Delta\frac{\lambda_{t,k,a}}{\lambda_{t,k,r}}$ in the occupation with a higher initial share of engagement in routine tasks. The results shown here parallel those presented in the theoretical model outlined by Autor, Levy, and Murnane (2003).

Further, let us examine how RBTC impacts self-selection by considering the simple case outlined by Autor and Handel (2013) where there are two occupations, *j* and *k*, with corresponding price vectors $\Lambda_{t,j} = \{\alpha_j, \lambda_{t,j,a}, \lambda_{t,j,r}\}$ and $\Lambda_{t,k} = \{\alpha_k, \lambda_{t,k,a}, \lambda_{t,k,r}\}$. For simplicity, assume that occupation *j* compensates only abstract tasks and occupation *k* only routine tasks implying that $\lambda_{t,j,a}, \lambda_{t,k,r} > 0$ and $\lambda_{t,j,r} = \lambda_{t,k,a} = 0$. Let us further assume that the population distribution of skill endowment for abstract and routine tasks takes a bivariate unit normal distribution:

$$\begin{bmatrix}\varepsilon_a \\ \varepsilon_r\end{bmatrix} \sim N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}1 & \sigma_{r,a} \\ \sigma_{a,r} & 1\end{bmatrix}\right) \qquad 5)$$

The difference in a worker's skill endowment is $v = (\varepsilon_a - \varepsilon_r)$ and the expected task endowment of workers who self-select into their occupation can be written as:

$$E[\varepsilon_a|i=j] = \frac{\lambda_{t,j,a}\lambda_{t,k,r}}{\sigma_v}\left(\frac{\lambda_{t,j,a}}{\lambda_{t,k,r}} - \rho\right)\left(\frac{\gamma(-z)}{\Gamma(z)}\right)$$

6)

$$E[\varepsilon_r|i=k] = \frac{\lambda_{t,j,a}\lambda_{t,k,r}}{\sigma_v}\left(\rho - \frac{\lambda_{t,k,r}}{\lambda_{t,j,a}}\right)\left(\frac{\gamma(-z)}{\Gamma(z)}\right)$$

The equations outlined in (6) characterize self-selection where $= (\alpha_k - \alpha_j)/\sigma_v$, $\rho = \sigma_{a,r}$, and $\gamma(z)/\Gamma(-z)$ is the inverse mills ratio. As noted by Autor and Handel (2013), a sufficient condition for positive self-selection is that $\rho \leq 0$ implying that the correlation between worker abilities is weakly negative. In their empirical analysis using self-reported measures of task engagement, the authors find sufficiently strong evidence supporting this condition and that self-selection occurs through comparative advantage.

Returning to the analysis of how RBTC impacts self-selection in the simple two occupation economy characterized by (6), consider the implications of increasing routine task productivity (e.g. $K_{t+1} > K_t$) on the expected task endowments. Recall that $\lambda_{t,j,r} = \lambda_{t,k,a} = 0$, implying that an increase in routine task productivity causes only a decrease in the premium for routine tasks $\frac{\partial \lambda_{t,k,r}}{\partial K_t} < 0$ and no change in the premium for abstract tasks $\frac{\partial \lambda_{t,j,a}}{\partial K_t} = 0$. Taking the partial derivative of the equations in (6) with respect to routine task productivity:

$$\frac{\partial E[\varepsilon_a|i=j]}{\partial K_t} = -\frac{\partial \lambda_{t,j,r}}{\partial K_t}\frac{\lambda_{t,j,a}}{\sigma_v}(\rho)\left(\frac{\gamma(-z)}{\Gamma(z)}\right)$$

$$\frac{\partial E[\varepsilon_r|i=k]}{\partial K_t} = \frac{\partial \lambda_{t,j,r}}{\partial K_t}\frac{\left(\lambda_{t,j,a}\rho - 2\lambda_{t,k,r}\right)}{\sigma_v}\left(\frac{\gamma(-z)}{\Gamma(z)}\right)$$

Since the routine task premium declines with an increase in factor productivity $\frac{\partial \lambda_{t,j,r}}{\partial K_t} < 0$ and empirical evidence suggests that $\rho \leq 0$, an increase in routine task productivity will cause the expected abstract task endowment of workers in occupation j to decrease (e.g. $\frac{\partial E[\varepsilon_a | i=j]}{\partial K_t} < 0$). As seen in the partial derivatives, RBTC will result in weaker self-selection into occupations with a high degree of engagement in routine tasks (e.g. occupations where the exogenous capital shock has the largest impact). Conversely, the same increase in routine task productivity will cause the expected routine task endowment of workers in occupation k to increase (e.g. $\frac{\partial E[\varepsilon_r | i=k]}{\partial K_t} < 0$) and thus result in stronger self-selection into that occupation. The intuition from this simple two occupation model carries over directly to the more general case of an economy with multiple occupations that utilize both abstract and routine tasks.[4]

2.C Theoretical Model: Empirical Implications

As shown in Section 2.B, the theoretical model outlined in Section 2.A has several important implications related to RBTC and the wage of workers. In particular, our model suggests that the routine task premium should decrease across all occupations while the premium for abstract tasks should increase over time. Further, the model predicts that the relative task premium will be more dramatically affected in occupations with a higher initial share of routine task engagement. In considering the self-selection of workers across occupations, the model also suggests that RBTC will create stronger self-selection of workers in routine-intensive occupations while simultaneously creating weaker self-selection in abstract-intensive occupations.

The empirical analysis proceeds by sequentially testing the following explicit theoretical implications of our model:

[4] Note that in the more general setting, a change in routine task productivity will create stronger (weaker) self-selection in occupations with an initial higher share of routine (abstract) task content.

1. $\forall t\ K_{t+1} > K_t \rightarrow \forall j\ \frac{\lambda_{t+1,j,a}}{\lambda_{t+1,j,r}} > \frac{\lambda_{t,j,a}}{\lambda_{t,j,r}}$; If the capital stock is exogenously increasing over time, the premium for abstract relative to routine tasks should increase across all occupations over time.
2. For two occupations $\{j, k\}$ such that $\frac{\beta_j}{(1-\beta_j)} < \frac{\beta_k}{(1-\beta_k)}$, $\forall t\ K_{t+1} > K_t \rightarrow \Delta \frac{\lambda_{t,j,a}}{\lambda_{t,j,r}} > \Delta \frac{\lambda_{t,k,a}}{\lambda_{t,k,r}}$; The premium for abstract relative should increase more in those occupations with a higher initial share of routine task engagement.
3. As illustrated in equation (6) and the subsequent discussion, changes in relative task premiums will lead to stronger selection into routine-intensive occupations and weaker selection into abstract-intensive occupations.

The empirical analysis contained in Section 4 sequentially examines each of the theoretical implications. In that analysis, I utilize a panel of occupational task content linked to panel data on workers and their employment arrangements. The advantage of this combined dataset is that I can exploit variation in task content across and, more importantly, within occupations over time. The within occupation variation in task content is particularly useful because it allows me to estimate changes in task premiums and the resulting self-selection of workers while control for unobserved skill through individual fixed-effects, e.g. $\phi_{i,a}$ and $\phi_{i,r}$ from equation (3). In addition, I can include occupation fixed-effects that control for output elasticity in the production of a final good, market demands, or a possible skill floor, e.g. ρ_j and α_j from equation (3). Fixed-effects estimation is critical because it allows me to link variation in task content with variation in wages in a manner that allows for a more robust test of the RBTC hypothesis.

3. Data Overview

The data used in this analysis combines a panel of individuals and their work activities with a panel of occupational task measures. The individual data comes from the 2004 and 2008 panels of the *Survey of Income Program Participation* (*SIPP*). The 2004 panel contains 12 waves of three months in length that stretch from October 2003 to December 2007 and the

2008 panel contains 16 waves that stretch from May 2008 to November 2013. A panel of occupational task measures was constructed exclusively from the survey data contained in 14 archived versions of the Occupational Information Network (*O*NET*) production database (i.e. analyst-updated data was dropped) released between April 2003 (*O*NET* 5) and July 2014 (*O*NET* 19). The panel of *O*NET* task measures was then linked to the *SIPP* panel by the occupation code of employed individuals.

The advantage of combining the *SIPP* with a synthetic panel of task measures is that I can exploit variation in task content within occupations over time and control for unobserved differences in individual productivity when assessing wage effects. Specifically, this combined framework allows me to track how the distribution of wages between and within occupations responds to short-run changes in task content. Focusing on short-run changes in task content ensures that identification is abject of occupation-specific institutional changes like the decline of unionization. Relative to using repeated cross-sections of the *Dictionary of Occupation Titles* (*DOT*), the *O*NET* panel has the advantage of being solely populated incumbent survey data and is thus less subject to mismeasurement.

3.A Data Overview: Panel of Occupational Tasks

The *O*NET* database was constructed as a replacement for the *DOT* (NRC 2010). Unlike the *DOT* which is populated by analyst-updated data, the *O*NET* was created with the goal of having the underlying measures collected from a survey of incumbent workers. The completed database was released in June 2002 (*O*NET* 4) with the initial measures having been populated by job analysts who assigned values to the *O*NET* survey questions by referencing the *DOT* releases from the 1980s. As a result, the initial release of the *O*NET* database was composed of an entirely new rating system but was populated with old analyst data using judgment-based methods.

Each year beginning in 2002, these initial analyst-updated fields were repopulated with new data from surveys administered to random samples of workers within specific occupations. On average, 110 different target occupations are revised in each of the 14 subsequent releases (*O*NET* 5 to 19) using newly collected incumbent survey data.[5] During a revision, each of the data fields in each occupation is replaced with the mean survey response from the newly collected data. As of the latest release (*O*NET* 19), there have been a total 589 7-digit *SOC* occupations that have been revised at least twice using surveys of incumbent worker. Changes in the content of these 589 occupations are driven solely by the survey responses of workers within those occupations and constitute the primary source of variation that I exploit in the empirical analysis. Although the *O*NET* has cautioned about using the data for time series analysis, the major concern relates to the intermingling of analyst and survey-updated data. As will be detailed below, I develop a careful methodology that relies only on survey-updated data and accounts for potential measurement error in the underlying data fields. Figure A.3 of the Technical Appendix provides a graphical depiction of the raw incumbent-updated sample sizes for all 7-Digit and 3-digit *SOC* occupations in *O*NET 19*.

The panel was constructed by first combining incumbent-updated measures from the work context and activity sections of each *O*NET* release. The value of each occupational task measure was linearly trended between the earliest and latest incumbent update. The occupations were aggregated from a 7-digit to a 3-digit *SOC* taxonomy using a rolling 3-year national employment weight constructed from the Occupational Employment Statistics. Estimates were conducted at a 3-digit *SOC* taxonomy to ensure a

[5] The *O*NET* selects occupations to be updated by considering a number of important factors that include but are not limited to the occupation's last update and a Department of Labor classification of a "demand-phase" occupation (Tippins & Hilton 2010, p. 5). The result is that occupations are sometimes selected for updates on the basis of relative employment size, demand, or changes in occupational content. Aggregating task measures from a 7-digit to a 3-digit *SOC* taxonomy using employment weights alleviates concern related to measurement error. The 7-digit occupations updated in each of the *O*NET* releases are distributed relatively evenly across the 3-digit *SOC* taxonomy. Assuming occupations are chosen for an update based on employment size and changes to content, the 3-digit aggregate measures will minimize measurement error and capture the underlying temporal variation.

sufficiently large and robust occupational sample size in both the *O*NET* and *SIPP*. The less detailed taxonomy also helps to assuage concerns about measurement error and selection.[6]

The abstract and routine task index in time period t for each occupation j is constructed from a weighted sum of the underlying task measures such that:

$$\bar{a}_{t,j} = \pi_t \left[\sum_{j \in J} W_j \sum_{k \in a} \left({LV_{t,j,k}}^{\frac{1}{3}} {IM_{t,j,k}}^{\frac{2}{3}} + CX_{t,j,k} \right) \right]$$
$$and \tag{7}$$
$$\bar{r}_{t,j} = \pi_t \left[\sum_{j \in J} W_j \sum_{k \in r} \left({LV_{t,j,k}}^{\frac{1}{3}} {IM_{t,j,k}}^{\frac{2}{3}} + CX_{t,j,k} \right) \right]$$

For each underlying task measure k, I utilize the level $LV_{t,j,k}$ and importance $IM_{t,j,k}$ scales if the measure is from the work activity category of the *O*NET* database but only the context $CX_{t,j,k}$ scale when the measure is from the context category. Following Blinder (2007) and Firpo et al. (2013), I assign a Cobb-Douglas weight of one third to level and two thirds to importance. After summing across all task measures $k \in (\cdot)$, each index is weighted by W_j which is equal to employment in occupation $j \in J$ (i.e. 6-digit *SOC* employment) relative to overall employment in occupation category J (i.e. 3-digit *SOC* employment).[7] The function π_t maps the raw employment-weighted task index to a percentile during each period. The resulting index values represent an occupation j's relative engagement in abstract or routine task content in period t. The underlying task variables within each index follow Autor and Handel (2013) and are detailed in Table 1. In relation to the theory from Section 2, the distributions of the two task indices $\bar{a}_{t,j}$ and $\bar{r}_{t,j}$

[6] Our use of employment weights also addresses problems related to changes in the *SOC* taxonomy during the analysis period. Specifically, I accomplish this by matching occupation codes in the *SIPP* to those in the *O*NET* panel at the 5,3, and 2-digit level respectively. Changes in the *SOC* taxonomy occur most frequently at the 6-digit level and, as a result, matching on higher level task measures provides an accurate imputation.

[7] As noted, the task indices were constructed with constant employment weights aggregated to a 3-digit *SOC* taxonomy. However, the estimates in the empirical analysis are robust to constructing the task indices using time variant employment weights or to aggregating to a 5,4, or 2-digit *SOC* taxonomy.

are assumed to be equivalent and act as suitable proxies for the true distribution of task content, i.e. the distribution of $\frac{\sum_{i\in j} a_{t,i}}{\sum_{i\in j} \ell^{*}_{t,j,i}}$ and $\frac{\sum_{i\in j} r_{t,i}}{\sum_{i\in j} \ell^{*}_{t,j,i}}$ which are measured in terms of efficiency units.[8]

[Insert Table 1]

At a given point in time, each of the indices can be thought of as relative measures of task engagement or, put differently, the occupational requirements necessary to produce a single unit of output. The advantage of using the panel of task measures, as opposed to a single cross-section, is that the task indices vary both across occupations and within occupations over time. In examining the distribution across our panel (i.e. comparing the 6-digit distribution in 2004 with 2014), there is a pronounced rightward shift that is statistically significant at the 98 percent level on a Kolmogorov–Smirnov test for both indices. Approximately half of the occupations exhibit movement upwards while the other half move downwards on the relative distribution. Asymptotic kernel density estimates of the 6-digit occupational distribution for the abstract and routine task index at the beginning and end of the panel are contained in Figures A.1 and A.2 of the Technical Appendix.[9]

In an effort to provide additional context around the changes to task content within the panel, the occupations with the highest growth/decline in task content were identified. The highest growth of engagement in abstract tasks from 2004 to 2014 occurred in *Fire Fighting and Prevention Workers* (92.3%) followed by *Nursing, Psychiatric, and Home Health Aides* (72.5%) and *Supervisors of Farming, Fishing, and Forestry Workers* (56.0%). The most significant decline in abstract task content was seen in *Animal Care and Service*

[8] The task indices can only be considered proxy variables for task content because the *O*NET* data is not scaled into any temporal unit of measure.

[9] Task measures were also constructed at the 2 and 5-digit *SOC* levels which had a larger and smaller sample size relative to the 3-digit measures used in the primary analysis. Estimation at these alternative aggregation levels was sufficiently convincing that sample selection or size was unlikely to have been a large factor driving the main result in Section 4.

Workers (57.1%) followed by *Other Protective Service Workers* (50.5%) and *Media and Communication Equipment Workers* (31.9%). The highest growth of routine task content occurred in *Entertainment Attendants and Related Workers* (70.3%) followed by *Supervisors of Building and Grounds Cleaning and Maintenance Workers* (64.8%) and *Nursing, Psychiatric, and Home Health Aides* (60.4%). The most significant decline in routine task content was seen in *Top Executives* (46.2%) followed by *Mathematical Science Occupations* (38.5%) and *Supervisors of Protective Service Workers* (35.2%).

3.B Data Overview: Panel of Individual Workers

The 2004 and 2008 *SIPP* panels were combined to create an unbalanced panel of approximately three million observations which was reduced to two million after sample restrictions. The *SIPP* is a household-based survey designed as a continuous representative series of national panels where the same individuals are interviewed over a multi-year period lasting approximately four years. The *SIPP* is the only available individual panel containing the necessary components to conduct an occupational analysis of prime-age workers. The *SIPP* has more detailed occupational codes, frequent interviews, and a larger sample than other comparable data sources.[10]

Descriptive statistics from the combined *SIPP* panels are presented in Table 2 where the sample has been limited to prime working age individuals between 25 and 55 years old who were not in the military. The estimation sample had a total of 83,018 individuals who were observed to be employed in the panel for an average 24.4 months,

[10] Compared to the Current Population Survey, its main advantage is the longitudinal nature that allows individuals and their job changes to be observed over time. Relative to the Panel Study of Income Dynamics, it provides a larger sample size, more frequent interviews and more detailed occupational codes. Although the level of detail of occupation codes is similar to that reported in the National Longitudinal Survey of Youth, the *SIPP* has much more frequent interviews and a larger sample with a more representative range of working age adults. In addition, the 2004 and 2008 *SIPP* panels were better aligned with the timing of the *O*NET* releases than National Longitudinal Survey of Youth.

totaling 2,023,051 observations.[11] On average, individuals reported working for 1.7 different employers and in 1.4 different occupations. As previously noted, the analysis considers occupations at a 3-digit *SOC* taxonomy. The average per period sample size in the *SIPP* across the 91 distinct 3-digit *SOC* occupations was 912.3 while the standard deviation was 1006.4 individuals. A histogram of the occupational sample size in the *SIPP* is shown graphically in Figure A.3 of the Technical Appendix where it is important to note that most occupations have a sufficiently large sample size.[12]

[Insert Table 2]

Employment information is reported in the *SIPP* under four distinct classifications: primary employment, secondary employment, primary self-employment, and secondary self-employment. All information for each of an individual's employment arraignments is recorded separately within each classification. Although an individual's occupation is recorded for secondary employment and self-employment, only the information recorded under an individual's primary employment arraignment was used for this analysis. Wages in the sample averaged 18.8 dollars per hour having been earned by workers who were on average 39.1 years of age with more working experience than formalized education. The sample was largely made up of workers with a high school degree or less and who reported their race as white and ethnicity as non-Hispanic. The descriptive statistics from Table 2 show that, although the *SIPP* does over sample low skill workers, the demographics of our estimation sample closely match those of the overall U.S. population.

---

[11] These figures vary based on the specification used in each part of the analysis. In addition to unreported occupational codes and other factors that cause observations to be omitted, our focus on wages limits the effective sample size to only those months where an individual reports employment.

[12] The three 3-digit occupations with a sample size of less than 30 individuals on average per period were distributed evenly across major 2-digit occupations. The 2 and 5-digit *SOC* levels were also used as a robustness check on our empirical findings at the 3-digit level. Estimation at these alternative aggregation levels was sufficiently convincing that sample selection occurring within the *SIPP* panel did not have an impact on the main result.

4. Empirical Analysis

This section details the results from an empirical analysis examining the implications of RBTC derived from the theoretical model presented in Section 2. The empirical analysis proceeds by sequentially testing the following explicit theoretical implications of the model: 1) the premium for abstract relative to routine tasks should increase over time for all occupations 2) the relative tasks premium should increase more in routine-intensive occupations 3) changes in the relative task premium should lead to stronger selection into routine-intensive occupations and weaker selection into abstract-intensive occupations. In the first subsection, I present evidence that the premium for routine task engagement is declining over time and more sharply in those occupations with an initially higher engagement in routine tasks. The second subsection uses the synthetic panel of occupational task content to examine within occupation wage effects and self-section.

4.A Empirical Analysis: Decline in the Routine Task Premium

Table 3 presents evidence in support of the RBTC narrative and that the routine task premium has decreased over time. The first panel, i.e. specifications 1 and 2, utilizes variation in task content within occupations over time and produces estimates using a two-step procedure motivated by Peri and Sparber (2009). The second and third panels, i.e. specifications 3 to 6, rely on cross-sectional variation in task content from individuals who change occupations within and outside of the same employer. As is discussed in more detail later, the results from Table 3 uniformly support the implications of the theoretical model pertaining to a declining price for routine task engagement.

Specifications 1 and 2 of Table 3 were estimated using a two-stage procedure where the second stage exploits temporal variation in task content within occupations and takes the form of a two-way fixed-effect model. In the first stage, a hedonic wage model was estimated for each year of data by regressing the log of hourly earnings on demographic characteristics and levels of human capital.[13] The unexplained wage variation captured by the residual was then predicted for each individual and averaged by period and occupation. In the second stage, the mean wage residual for each occupation was regressed on the indices for abstract and routine task content using a model with occupation and time fixed-effects.[14] The two-stage strategy estimates changes to task premiums/penalties only after controlling for supply-side characteristics. The first specification suggests an increase of one percentile in the abstract task index would yield a 34.4 percentage point increase in hourly wages. In contrast, an increase of one percentile in the routine task index would yield only a 3.2 percentage point increase in hourly earnings. The second specification which interacts the routine task index with a linear time trend reveals that the premium is not only declining but eventually inverting over time.

[Insert Table 3]

The distribution of residual wages, predicted by the first-stage hedonic wage model from specifications 1 and 3 of Table 3, is depicted graphically in Figure 1 where the data has been smoothed using kernel density estimation. The first and second panel plot the distribution of residual wages in 2004 (*O*NET* 6) and 2013 (*O*NET* 18) for those individuals employed in occupations that fell into the first and fifth quintile of the routine

---

[13] The first stage hedonic wage model was applied as a rolling cross-sectional regression estimated each time period and took the form $w_{t,i} = \gamma_t + X_{t,i}\beta_X + H_{t,i}\beta_H + \mu_{t,i}$ where $w_{i,t}$is the log of hourly wages, γ_t is a time period fixed-effect, $X_{t,i}$ is a matrix of demographic characteristics (i.e. age, number of children, marital status, gender, race, and ethnicity), $H_{t,i}$ is a matrix of human capital measures (i.e. job tenure, job tenure squared, and educational attainment), and $\mu_{t,i}$ is the individual wage residual.

[14] The second stage fixed-effect model was applied at the occupation level and took the form $\widehat{w}_{t,j} = \gamma_t + \alpha_j + \lambda_a \bar{a}_{t,j} + \lambda_r \bar{r}_{t,j} + \mu_{t,j}$ where $\widehat{w}_{t,j}$ is the mean standardized wage residual from the first stage hedonic, γ_t is a time period fixed-effect, α_j is set of 91 occupation dummy variables, $\bar{a}_{t,j}$ is the percentile rank of occupation j's abstract task engagement in period t, $\bar{r}_{t,j}$ is the percentile rank of occupation j's routine task engagement in period t, and $\mu_{t,j}$ is the residual error term. The second specification interacts the routine task index with a linear time trend such that $(\bar{r}_{t,j} \cdot t)$ and where the coefficient represents the economy-wide change in the routine task premium.

task distribution in 2004. In comparing these two panels, the residual wage distribution shifts rightward for the first quintile but leftward for the fifth and both shifts were statistically significant on a Kolmogorov–Smirnov at the 99 percent level. The third and fourth panels plot the distribution of residual wages for those employed in occupations from the first and fifth quintile of the abstract task distribution. These distributions exhibit the opposite pattern as that observed in the routine task distribution, namely the first quintile shifts leftward overtime while the fifth quintile shifts rightward. As before, both of these shifts were statistically significant on a Kolmogorov–Smirnov at the 99 percent level.

[Insert Figure 1]

The results shown for specifications 3 through 6 of Table 3 are comparable to specifications 1 and 2 in terms of statistical significance and magnitude but were estimated using an alternative procedure. Rather than controlling for supply-side characteristics using a two-stage procedure, I control for these factors by exploiting variation in task content and wages from employer-to-employer occupational transitions. Specifically, I estimate changes in the task premium by regressing the change in log wages from an employer-to-employer transition where the individual also changed occupations on the cross-sectional variation in task content across the individual's old and new occupation.[15] Specification 5 and 6 apply the same model as specifications 3 and 4 but control for firm heterogeneity by restricting the sample to those individuals who changed occupations but remained with the same employer. The results from this alternative estimation strategy also support the implications from the theoretical model that suggests the premium for routine task engagement should be declining over time.

[15] In the third through sixth specifications of Table 3, the fixed-effect model takes the form $w_{t,j,i} = \gamma_t + \delta_i + \lambda_a \bar{a}_{t,j} + \lambda_r \bar{r}_{t,j} + \mu_{t,j,i}$ where γ_t is a time period fixed-effect, δ_i is a person fixed-effect, $\bar{a}_{t,j}$ is the percentile rank of occupation j's abstract task engagement in period t, $\bar{r}_{t,j}$ is the percentile rank of occupation j's routine task engagement in period t, and $\mu_{t,j,i}$ is the residual error term. The fourth and sixth specification in Table 3, interacts the routine task index with a linear time trend such that $(\bar{r}_{t,j} \cdot t)$ and where the coefficient represents the economy-wide change in the routine task premium.

Figure 2 plots the coefficient on routine task content from the most restrictive set of estimates in Table 3, namely specifications 5 and 6. As suggested by the theoretical model, those occupations with an initially higher level of routine task content received a higher relative wage premium in the initial period. As confirmed by specifications 2, 4, and 6 of Table 3, the premium paid for engagement in routine tasks has declined since 2003. Moe specifically, it has declined more dramatically in routine-intensive occupations. The graphical depiction of the coefficient from Table 3 emphasizes the fact that these estimates support the predictions of the theoretical model.

[Insert Figure 2]

4.B Empirical Analysis: Wage Effects and Self-Selection

This subsection presents estimates that rely on the panel of occupational task content for identification and include a series of individual, occupation, and job-spell fixed-effects. These estimates quantify the impact of changes in task content on an individual's wage. The results strongly support the theoretical prediction that RBTC is driving self-selection. In particular, self-selection has become stronger in routine-intensive occupations but weaker in abstract-intensive occupations. The estimates include only task indices for abstract and routine but estimates that include an additional index for non-routine task engagement, are included in Tables A.1, A.2, and A.3 of the Technical Appendix. In the fifth section, additional steps are taken towards extending these estimated towards causal inference where instrumental variables are applied using two-stage least squares.

I estimate the effect of changes in occupational task content on wages by applying the following model with two-way fixed-effects:

$$w_{t,j,i} = \gamma_t + \delta_i + \lambda_a \bar{a}_{t,j} + \lambda_r \bar{r}_{t,j} + \mu_{t,j,i} \tag{8}$$

The wage of individual i in occupation j during period t is estimated such that γ_t is a time period fixed-effect, δ_i is an individual fixed-effect, and $\mu_{t,j,i}$ is the residual error term.

The two variables of interest are $\bar{a}_{t,j}$ and $\bar{r}_{t,j}$ are the percentile rank of occupation j's abstract and routine task content in period t. The results from applying equation (8) to the data are contained in the first specification of Table 4 where the two task indices are found to be statistically insignificant. The individual fixed-effect controls for unobserved individual productivity differences but, by construction, allows for both within and between variation in the two task indices. The variation in the two task indices comes from changes in occupational task content if an individual stays in the same occupation throughout the panel. If an individual changes occupations, however, the task indices will also pick-up variation from cross-sectional differences in task content between occupations.

In an effort to control for additional unobservable heterogeneity and limit the variation to changes in task content within occupations, I sequentially introduce an increasingly restrictive set of fixed-effects. The second specification in Table 4 appends equation (8) with α_j a set of 91 occupation dummies. The third specification replaces δ_i and α_j with an occupation-spell fixed-effect $(\delta_i \cdot \alpha_j)$ captured by 119,834 unique dummy variables for each occupation that each individual holds in the panel. Annotating a set of employer dummies with the variable η_e, the fourth specification replaces δ_i and α_j with a job-spell fixed-effect $(\delta_i \cdot \eta_e)$ captured by 140,325 unique dummy variables for each job (employer) that each individual holds in the panel. The last and most comprehensive specification replaces δ_i and α_j with an occupation-job-spell fixed-effect $(\delta_i \cdot \alpha_j \cdot \eta_e)$ captured by 146,383 unique dummy variables for each occupation that each individual holds with each employer in the panel.

[Insert Table 4]

Imposing increasingly restrictive fixed-effects across each specification in Table 4 helps to pinpoint the source of variation driving each estimate. The fifth and most restrictive specification controls for unobserved differences in an individual's productivity within a specific occupation at a specific firm. The inclusion of this fixed-effect ensures that the variation captured in the coefficient estimates comes exclusively from changes in task content that occur as an individual's job function evolves over time. In particular, identification comes exclusively from changes in the task content of occupations over time and not from mobility of individuals across occupations or jobs. An increase of one percentile in the abstract task index corresponds with a 5.9 percentage point increase in an individual's hourly wages. In contrast, an increase of one percentile in the routine task index yields an 8 percentage point decrease in an individual's hourly earnings.

The theoretical model from Section 2 suggests that, as the factor productivity of routine tasks increases, the mean level of routine task engagement declines because capital substitutes for routine tasks (i.e. workers using the new technology become more productive at them). As discussed, we expect these changes to occur more sharply in those occupations that are initially more routine-intensive. Using a framework outlined by Autor and Handel (2013), implications derived in Section 2 suggest that technology should cause self-selection to become stronger in routine-intensive occupations and weaker in abstract-intensive occupations. Although the results shown in Tale 4 preliminary support RBTC's implications about self-selection, they are explored in substantially more detail in the context of Table 5.

Table 5 reports a correlate for Table 4 with three additional terms added, namely a linear time trend and its interaction with the two task indices. In examining the coefficient estimates across each specification, the results present compelling evidence that self-selection has increased substantially over time. In particular, our theoretical model of RBTC suggests that routine task content should be decreasing across occupations as capital (e.g. factor productivity) increases. Since the task indices represent a proxy for the occupational average of task content per worker, we can interpret a decrease of one percentile in routine task content as increasing wages by 6.8 percentage points and this coefficient increasing by 0.2 percentage points across the 14 years in the panel. The higher wages associated with declining routine task content provide additional evidence that supports the idea that self-selection increases in routine-intensive occupation. Similarly supported by the theory, an increase in abstract task content moves wages in the opposite direction and trends differently over time.

[Insert Table 5]

Table 6 reports the results from applying the between effects correlate of equation (8) to the panel. In the current context, the between effects estimator utilizes the cross-sectional information in the data by averaging the task variables across all time periods for the requisite level of fixed effects. As before, the fifth and final specification should be considered the most comprehensive as it captures unobserved heterogeneity for each occupation that each individual holds at each of her employers. An increase of one percentile in the abstract task index across occupations corresponds with a 96.9 percentage point increase in hourly wages. In contrast, an increase of one percentile in the routine task index across occupations corresponds with a 15.7 percentage point increase in an individual's hourly earnings. As with the fixed-effects estimates, the cross-sectional estimates are consistent with the RBTC narrative and discussion from the theory section, namely that routine-intensive occupations are paid a relatively higher way and that task premiums must be positive.

[Insert Table 6]

5. Identification of Causality

The underlying measures of task content, used to construct the synthetic panel of occupational task content, are derived from *O*NET* survey data. As such, movement in the task measures is not only driven by changes to the production process (e.g. new technology) but also by flows of workers into and out of different, i.e. a compositional change in the workers completing the surveys. If individuals sort across occupations based on unobservable characteristics like heterogeneous time-variant expectations about changes to task content and technology, the estimates of between and within wage effects (Section 4.A) would suffer from an endogeneity problem. Specifically, we should expect that the estimates will underestimate the effect of changes to task content across occupations and within occupations over time. I make progress towards identifying the causal impact of changes to task content on wages by constructing instrumental variables for the two task indices and applying two-stage least squares estimation.

I construct an index that proxies for economy-wide changes to the production process and then use that index as an instrument for changes in occupational task content. The logic and structure of the instrument is highly motivated by the enclave instruments used frequently in the immigration literature and originally proposed in Altonji and Card (1991) and Card (2001). According to the theoretical model (Section 2), those occupations with an initially higher level of abstract and routine task content will be affected more by aggregate changes to task content. The logic of the instrument is based on the idea that the distribution of task content for an occupation *j* in an initial time period $t_0$will impact the extent that economy-wide changes to technology affect task content within that occupation. The instrument is, by construction, uncorrelated with idiosyncratic changes in task content at the occupation level which are driven by things like the sorting of workers across occupation based on expectations about the impact and trajectory of technological change.

The abstract $\bar{a}_{t,j}$ and routine $\bar{r}_{t,j}$ measures of task content for occupation *j* in time period *t*, detailed in (8) of Section 3, were instrumented using the following variables:

$$\widehat{(\bar{a}_{t,J})} = \sum_{j \in J} W_j \sum_{k \in a} \left[LV_{t,j0,k}{}^{\frac{1}{3}} IM_{t,j0,k}{}^{\frac{2}{3}} \frac{LV_{t0,j,k}{}^{\frac{1}{3}} IM_{t0,j,k}{}^{\frac{2}{3}}}{LV_{t0,j0,k}{}^{\frac{1}{3}} IM_{t0,j0,k}{}^{\frac{2}{3}}} + CX_{t,j0,k} \frac{CX_{t0,j,k}}{CX_{t0,j0,k}} \right]$$

$$and \tag{9}$$

$$\widehat{(\bar{r}_{t,J})} = \sum_{j \in J} W_j \sum_{k \in r} \left[LV_{t,j0,k}{}^{\frac{1}{3}} IM_{t,j0,k}{}^{\frac{2}{3}} \frac{LV_{t0,j,k}{}^{\frac{1}{3}} IM_{t0,j,k}{}^{\frac{2}{3}}}{LV_{t0,j0,k}{}^{\frac{1}{3}} IM_{t0,j0,k}{}^{\frac{2}{3}}} + CX_{t,j0,k} \frac{CX_{t0,j,k}}{CX_{t0,j0,k}} \right]$$

A Cobb-Douglas weighting of one third was applied to the level ($LV_{t,j0,k}$) and two thirds to the importance ($IM_{t,j0,k}$) scales of task measure *k* across all occupations j_0 at time period *t*. This economy-wide measure of task content at time period *t* was then adjusted by a constant relative share factor. The share factor was constructed by dividing task content for occupation *j* by the level across all occupations j_0 in time period t_0.[16] Abject the Cobb-Douglas weighting scheme, the same process was carried out for task measures reported

[16] Task measures occurring during the base time period t_0 were drawn from *O*NET 4* where all task measures were populated by analysts using the 1984 *DOT* as the primary reference source.

using the context scale. As seen in (9), each index is weighted by W_j which is equal to employment in occupation $j \in J$ (i.e. 6-digit SOC employment) relative to overall employment in occupation category J (i.e. 3-digit SOC employment).

Table 7 contains second stage estimates of wage effects within occupations obtained using two-stage least squares where an increasingly restrictive set of fixed-effects are introduced across each specification.[17] As mentioned, these estimates control for possible time variant selection based on expectations about the future of the labor market. If we believe that individuals have expectations about how task engagement might evolve over time, instrumental variable estimation makes progress towards identifying the causal impact of changes to task content on wages. In terms of statistical significance, the estimates correspond with our hypothesis that changes to task content affect the wages of incumbent workers. Although larger in terms of magnitude, these estimates are similar and move in the hypothesized direction when considered relative to the estimates in Table 4.

[Insert Table 7]

As a correlate for Table 6, Table 8 reports the results from applying the between using two-stage lease squares estimation. effects correlate of equation (8) to the panel.[18] As discussed previously, the between effects estimator utilizes the cross-sectional information in the data by averaging the task variables across all time periods for the requisite level of fixed effects. In terms of statistical significance, the estimates correspond with our hypothesis that changes to task content affect the wages of incumbent workers. As was the case I comparing Table 7 with Table 4, the estimates in Table 8 are similar and move in the hypothesized direction when considered relative to the estimates in Table 6.

[Insert Table 8]

---

[17] The F-Statistic and R-Squared from the first stage estimates illustrate that our instrument is highly correlated with the variable of interest. Tables that detail these results are provided in Appendix Table A.4.

[18] The F-Statistic and R-Squared from the first stage estimates illustrate that our instrument is highly correlated with the variable of interest. Tables that detail these results are provided in Appendix Table A.4.

6. Conclusion

In this study, I develop a theoretical model of occupational production that is a natural extension to Autor and Handel (2013) and use the model to derive relevant implications related to RBTC and self-selection. As is consistent with the literature on RBTC, the model predicts the following empirically testable implications: 1) the premium for abstract relative to routine tasks should unequivocally increase over time for all occupations 2) the relative tasks premium should increase more in routine-intensive occupations 3) changes in the relative task premium should lead to stronger selection into routine-intensive occupations and weaker selection into abstract-intensive occupations.

The predictions of the RBTC hypothesis are tested empirically by directly examining wage and employment dynamics using a combined panel of occupational task content and individual workers. The construction of a novel panel on occupational task content allows for further testing of the RBTC hypothesis by examining the consequent sorting of workers based on comparative advantage as well as changes to relative wages over time. In the estimates, I also control for time invariant factors like unobserved heterogeneity through estimation using individual, occupation, and job-spell fixed-effects. I find new and compelling evidence in support of the RBTC hypothesis. More specifically, I find evidence that the relative premium for routine tasks has declined most rapidly in routine-intensive occupations and that self-selection has increased sharply over time. The estimates obtained from this analysis provide a new perspective on RBTC and offer additional insight into how labor market dynamics are impacted by technological change. In addition, the methodology used to construct the panel from archived releases of the *O*NET* could be used for a variety of additional applications.

References

Acemoglu, Daron & David Autor. 2011. Skills, Tasks and Technologies: Implications for Employment and Earnings," Handbook of Labor Economics, Elsevier.

Altonji, Joseph G. & David E. Card. 1991. The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives. NBER Chapters, in: Immigration, Trade, and the Labor Market, pages 201-234 National Bureau of Economic Research, Inc.

Autor, David H. & Lawrence F. Katz & Alan B. Krueger. 1998. Computing Inequality: Have Computers Changed The Labor Market? The Quarterly Journal of Economics. MIT Press, vol. 113(4) (November): 1169-1213

Autor, David H. & Frank Levy & Richard J. Murnane, 2003. The Skill Content Of Recent Technological Change: An Empirical Exploration. The Quarterly Journal of Economics. MIT Press, vol. 118(4). (November):1279-1333

Autor, David H. & Lawrence F. Katz & Melissa S. Kearney. 2005. Trends in U.S. Wage Inequality: Re-Assessing the Revisionists. NBER Working Papers 11627, National Bureau of Economic Research, Inc.

Autor, David H. 2013. The "Task Approach" to Labor Markets: An Overview. Journal of Labour Market Research.

Autor, David H. & Michael J. Handel. 2013. Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. Journal of Labor Economics. University of Chicago Press, vol. 31(S1): S59 - S96.

Autor, David H. 2014. Polanyi's Paradox and the Shape of Employment Growth. NBER Working Papers 20485, National Bureau of Economic Research, Inc.

Altonji, Joseph G. & Lisa B. Kahn & Jamin D. Speer. 2014. Trends in Earnings Differentials across College Majors and the Changing Task Composition of Jobs. American Economic Review. American Economic Association, vol. 104(5) (May): 387-93

Blinder, Alan S. 2007. How Many Jobs Might be Offshorable? Center for Economic Policy Studies Working Discussion Paper no. 142, Princeton University.

Böhm, Michael. 2015. The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change. University of Bonn Working Paper. September

Card, David. 1991. Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration. *Journal of Labor Economics.* University of Chicago Press, vol. 19 (1), 22-64.

Cortes, Guido Matias & Nir Jaimovich & Christopher J. Nekarda & Henry E. Siu. 2014. The Micro and Macro of Disappearing Routine Jobs: A Flows Approach. NBER Working Papers 20307, National Bureau of Economic Research, Inc.

Cortes, Guido Matias. 2016 (Forthcoming). Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data. Journal of Labor Economics. University of Chicago Press, vol. 34 (1)

Firpo, Sergio & Nicole M. Fortin & Thomas Lemieux. 2013. Occupational Tasks and Changes in the Wage Structure. University of British Columbia Working Paper.

Gathmann, Christina & Uta Schönberg. 2010. How General Is Human Capital? A Task-Based Approach. Journal of Labor Economics. University of Chicago Press, vol. 28(1): 1-49, 01.

Jensen, Bradford J. & Lori G. Kletzer. 2010. Measuring Tradable Services and the Task Content of Offshorable Services Jobs. NBER Chapters, in: Labor in the New Economy, National Bureau of Economic Research: 309-335

Katz, Lawrence F. & Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors. The Quarterly Journal of Economics. MIT Press, vol. 107(1) (February): 35-78

National Research Council (NRC). 2010. "A Database for a Changing Economy: Review of the Occupational Information Network (*O*NET*)". Panel to Review the Occupational Information Network (*O*NET*). Nancy T. Tippins and Margaret L. Hilton, editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Peri, Giovanni & Chad Sparber. 2009. Task Specialization, Immigration, and Wages. American Economic Journal: Applied Economics, American Economic Association, vol. 1(3), pages 135-69, July.

Spitz-Oener, Alexandra. 2006. "Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure". Journal of Labor Economics. University of Chicago Press, vol. 24 (2)

Yamaguchi, Shintaro. 2011. Tasks and Heterogeneous Human Capital. Journal of Labor Economics. University of Chicago Press, vol. 30(1): 1 – 53

Tables and Figures

Table 1: Composition of the Abstract and Routine Task Indices

Task Index	*O*Net* Database	
	Task Measure	Scale
Abstract	Analyzing Data or Information	Work Activities
	Thinking Creatively	Work Activities
	Interpreting the Meaning of Information for Others	Work Activities
	Establishing and Maintaining Interpersonal Relationships	Work Activities
	Guiding, Directing, and Motivating Subordinates	Work Activities
	Coaching and Developing Others	Work Activities
Routine	Importance of Being Exact or Accurate	Work Context
	Importance of Repeating Same Tasks	Work Context
	Structured versus Unstructured Work	Work Context
	Controlling Machines and Processes	Work Activities
	Spend Time Making Repetitive Motions	Work Context
	Pace Determined by Speed of Equipment	Work Context

Table 2: Descriptive Statistics from *SIPP* Estimation Sample

Survey Characteristics	Observations	Individuals	Avg. Survey Response	Avg. Number of Jobs	Avg. Number of Occupations
	2,023,051	83,018	24.37	1.69	1.44
Workplace Characteristics	Hourly Wage	Age	Years of Education	Experience	Hours
	18.76	39.12	13.69	19.85	34.90
	(36.45)	(9.14)	(2.59)	(43.24)	(13.55)
Education	Less than HS	HS	Some College	College	Post-College
	8.66%	41.74%	20.37%	19.60%	9.63%
Demographics	White	Black	Asian	Other	Hispanic
	79.50%	12.38%	4.30%	3.82%	12.48%
Gender	Male			Female	
	50.11%			49.89%	

Table 3: Evidence of a Declining Price for Routine Task Engagement

Dependent: Log Hourly Wage		Task Content Within Occupations		Task Content Across Occupations, One Employer		Task Content Across Occupations	
		(1)	(2)	(3)	(4)	(5)	(6)
Abstract		0.344***	0.334***	0.14***	0.142***	0.238***	0.24***
		(0.001)	(0.001)	(0.016)	(0.017)	(0.029)	(0.029)
Routine		0.032***	0.064***	0.048**	0.29***	0.032	0.131***
		(0.002)	(0.002)	(0.023)	(0.058)	(0.026)	(0.051)
Routine x Year			-0.006***		-0.026***		-0.008***
			(0.000)		(0.006)		(0.003)
Fixed-Effects	Individual			Yes	Yes	Yes	Yes
	Time	Yes	Yes	Yes	Yes	Yes	Yes
	Occupation	Yes	Yes				
Observations		356,000		8,956		49,727	
Groups		91 Occ.		4,007 Ind.		18,377 Ind.	

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on 91 of the 3-digit *SOC* occupations reported in the *SIPP*.

Figure 1: Kernel Density Estimates of Wage Residual by Task Distribution, 2004-13

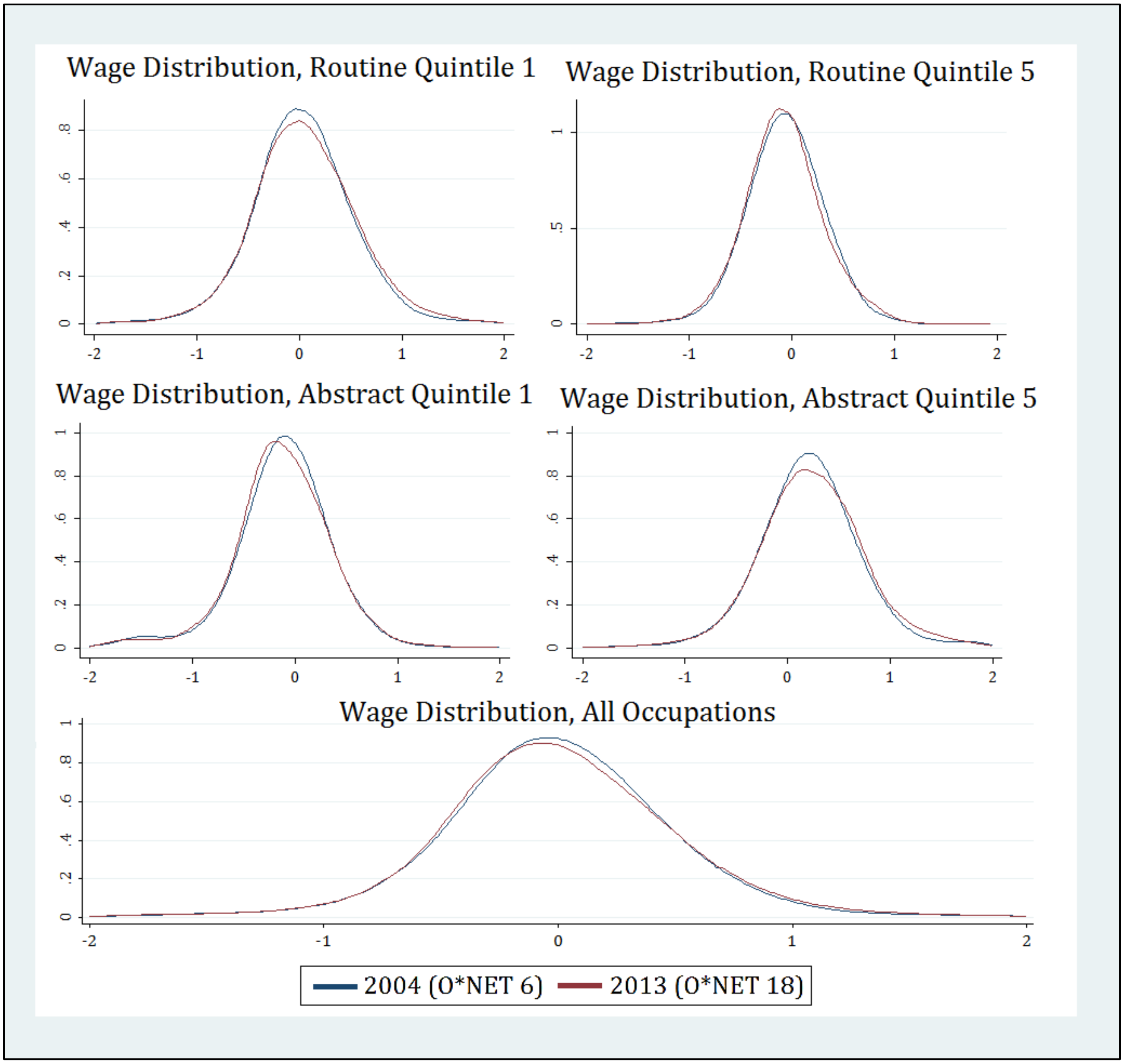

Figure 2: Graphical Evidence of a Declining Price for Routine Task Engagement

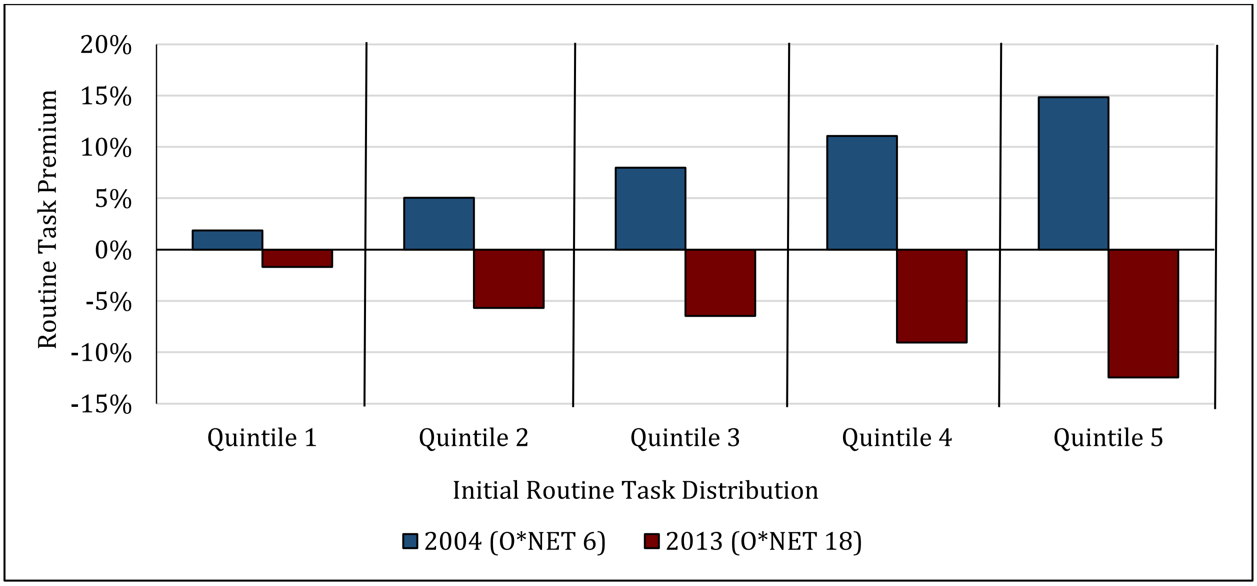

Table 4: Fixed-Effects Regression of Log Wages on Task Indices

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		0.248 (0.002)	-0.001 (0.007)	0.02** (0.009)	0.145*** (0.004)	0.061*** (0.009)
Routine		0.031 (0.003)	-0.037*** (0.008)	-0.045*** (0.01)	-0.033*** (0.005)	-0.083*** (0.01)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Table 5: Fixed-Effects Regression of Log Wages on Task Indices Interacted with Year

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		0.182*** (0.004)	0.003 (0.007)	0.040*** (0.009)	0.123*** (0.005)	0.072*** (0.009)
Routine		0.039*** (0.004)	-0.024*** (0.009)	-0.012 (0.011)	-0.016*** (0.006)	-0.068*** (0.011)
Year		0.094*** (0.006)	0.114*** (0.006)	0.116*** (0.005)	0.101*** (0.005)	0.113*** (0.005)
Abstract x Year		0.010*** (0.000)	0.004*** (0.001)	-0.000 (0.001)	0.008*** (0.001)	0.001 (0.001)
Routine x Year		-0.000 (0.001)	-0.003*** (0.001)	-0.005*** (0.001)	-0.002*** (0.001)	-0.002*** (0.001)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Table 6: Cross-Sectional Regression of Log Wages on Task Indices

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		1.078*** (0.009)	-0.053* (0.031)	0.941*** (0.007)	0.992*** (0.007)	0.969*** (0.007)
Routine		0.197*** (0.009)	0.163*** (0.036)	0.138*** (0.008)	0.168*** (0.007)	0.157*** (0.007)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Table 7: Instrumental Variable Fixed-Effects Regression of Log Wages on Task Indices

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		0.429***	0.0721	0.236**	0.285***	1.439***
		(0.00501)	(0.208)	(0.120)	(0.00953)	(0.165)
Routine		0.252***	-0.965**	-0.976***	0.163***	-2.653***
		(0.00544)	(0.446)	(0.211)	(0.0116)	(0.267)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Table 8: Instrumental Variable Cross-Sectional Regression of Log Wages on Task Indices

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		1.504***	-0.093	1.325***	1.406***	1.376***
		(0.017)	(0.726)	(0.014)	(0.014)	(0.013)
Routine		0.672***	0.073	0.572***	0.631***	0.615***
		(0.017)	(1.565)	(0.014)	(0.014)	(0.013)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Technical Appendix

Figure A.1: Kernel Density Estimates of Abstract Task Index, 2004-14

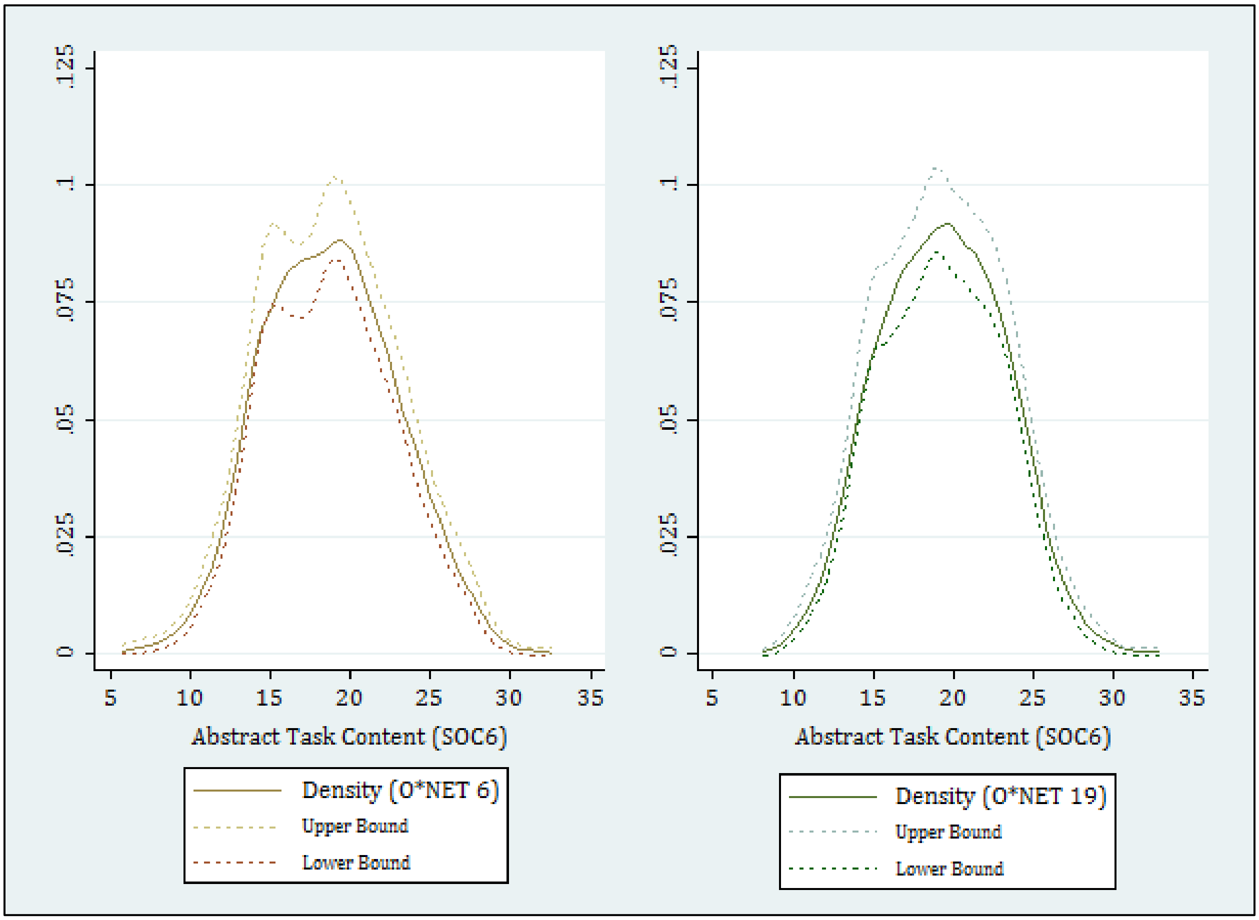

Figure A.2: Kernel Density Estimates of Routine Task Index, 2004-14

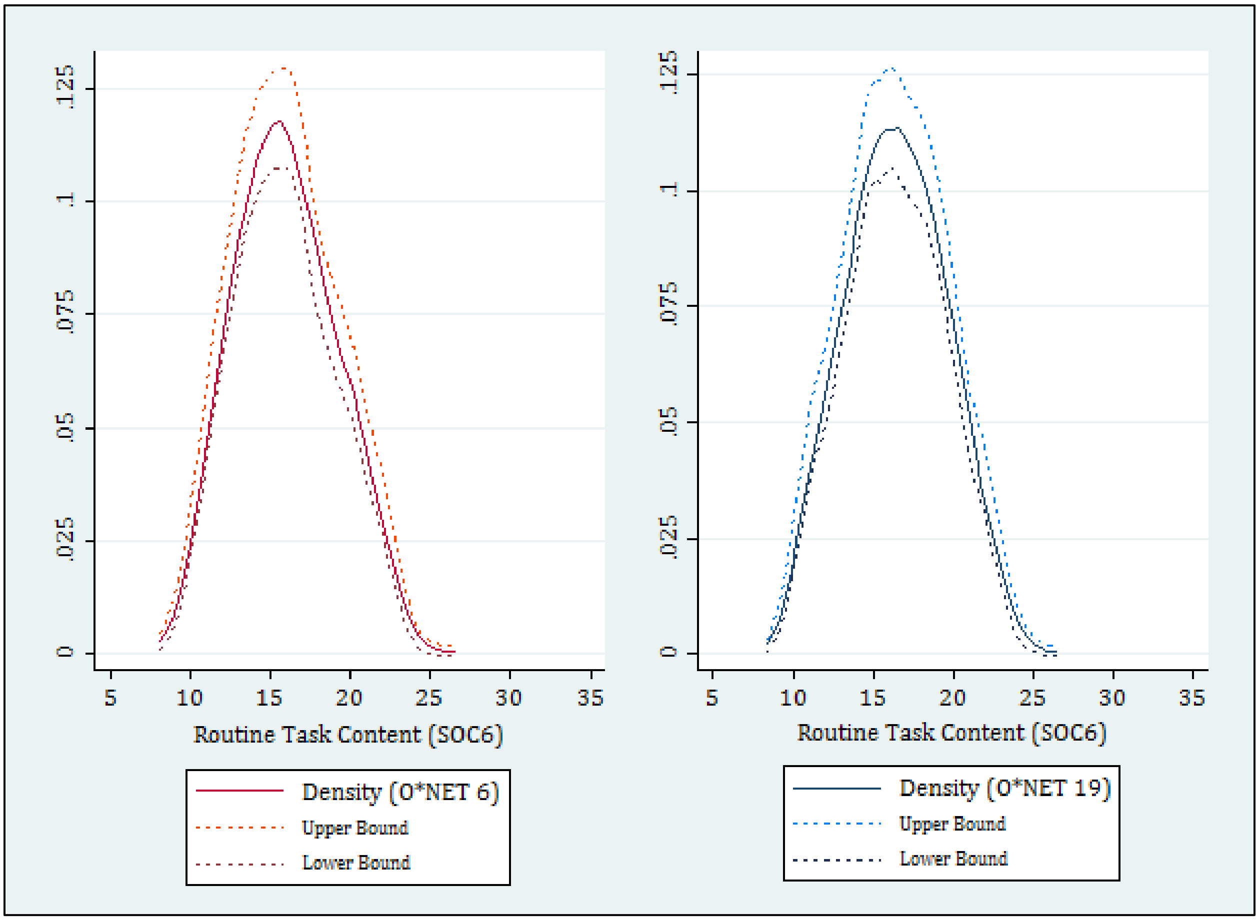

Figure A.3: Histogram of Incumbent-Updated Survey Data in *O*NET* 19

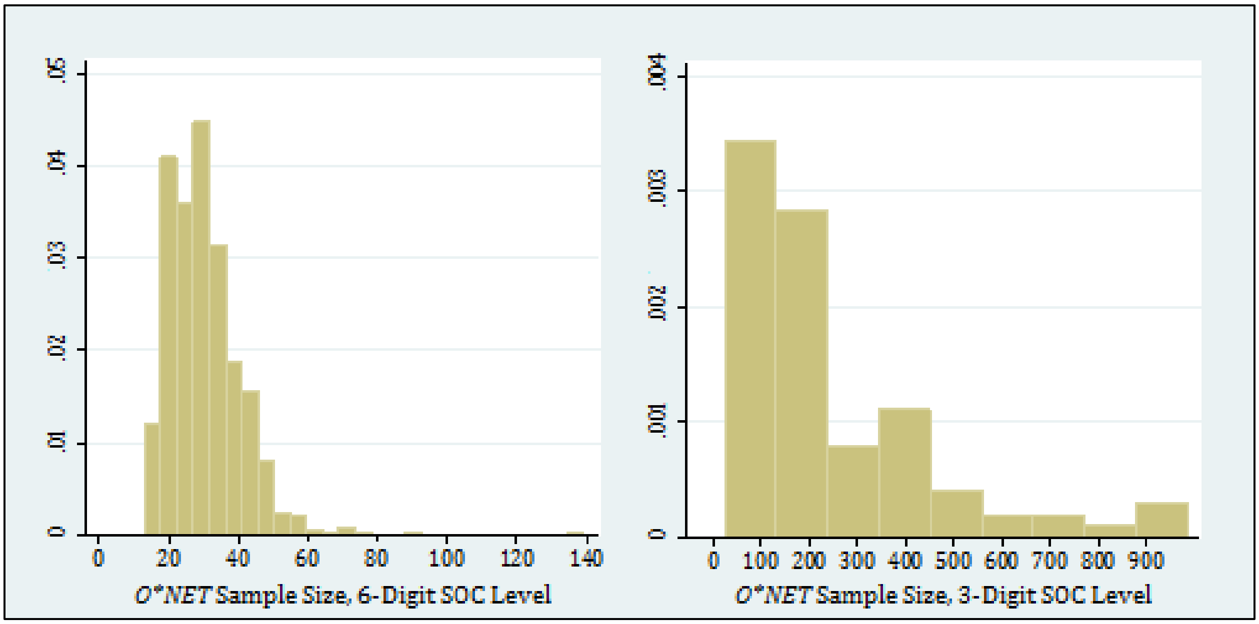

Figure A.4: Histogram of Average Occupational Sample Size in the *SIPP* Estimation Sample

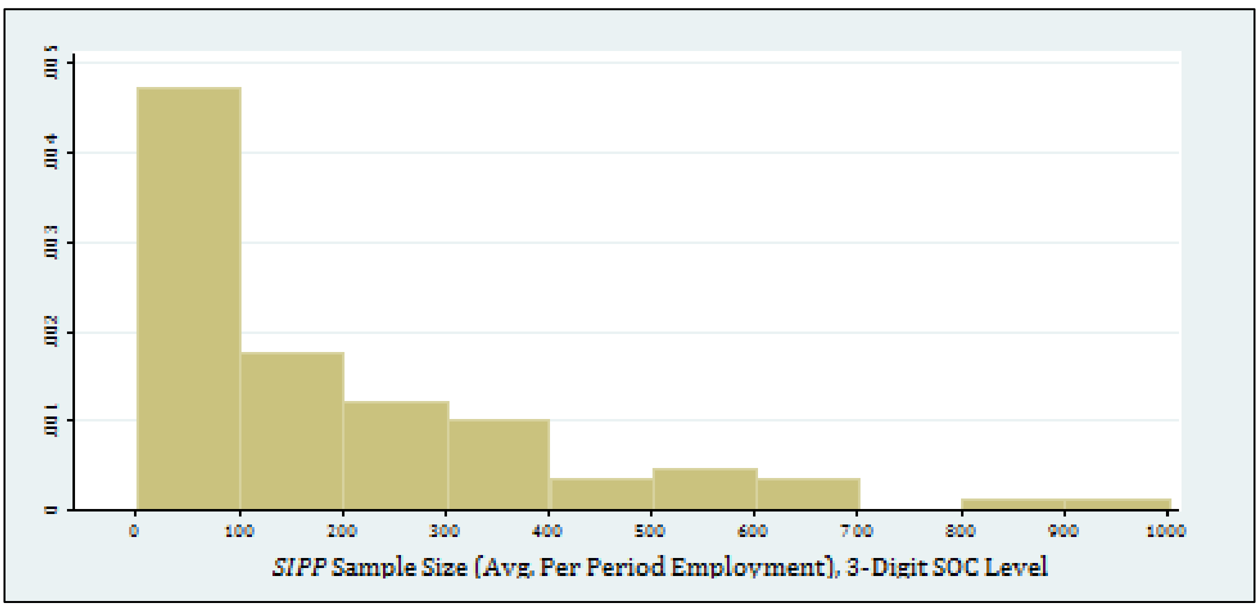

Table A.1: Three Factor Estimates of Fixed-Effects Regression of Log Wages on Task Indices

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		0.229***	0.029	0.094***	0.114***	0.111***
		(0.008)	(0.023)	(0.029)	(0.013)	(0.03)
Non-Routine Manual		-0.077***	-0.023	-0.157***	-0.038**	-0.129***
		(0.01)	(0.039)	(0.046)	(0.019)	(0.047)
Routine		0.06***	-0.094***	-0.103***	-0.016	-0.127***
		(0.01)	(0.025)	(0.031)	(0.018)	(0.032)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Table A.2: Fixed-Effects Regression of Log Wages on Task Indices Interacted with Year

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		0.181***	0.011	0.056***	0.121***	0.081***
		(0.004)	(0.008)	(0.010)	(0.005)	(0.010)
Routine		0.076***	-0.015*	-0.003	0.020***	-0.055***
		(0.005)	(0.009)	(0.011)	(0.007)	(0.011)
Non-Routine Manual		-0.066***	-0.043***	-0.069***	-0.076***	-0.057***
		(0.005)	(0.013)	(0.016)	(0.007)	(0.015)
Year		0.094***	0.114***	0.116***	0.101***	0.112***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Abstract x Year		0.010***	0.004***	0.000	0.008***	0.001
		(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Routine x Year		-0.001	-0.004***	-0.005***	-0.001	-0.004***
		(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Non-Routine Manual x Year		0.001	0.001**	0.000	-0.001	0.002***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Table A.3: Three Factor Estimates of Cross-Sectional Regression of Log Wages on Task Indices

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
Abstract		1.068***	0.018	0.966***	1.003***	0.985***
		(0.01)	(0.038)	(0.009)	(0.008)	(0.008)
Non-Routine Manual		-0.086***	-0.115*	-0.092***	-0.082***	-0.083***
		(0.011)	(0.069)	(0.009)	(0.009)	(0.009)
Routine		0.242***	0.131***	0.202***	0.217***	0.209***
		(0.013)	(0.042)	(0.012)	(0.011)	(0.011)
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.

Table A.4: First Stage Regression of Task Indices on Instrumental Variables

Dependent: Log Hourly Wage		(1)	(2)	(3)	(4)	(5)
				1st Stage: Abstract		
Abstract		0.031***	-0.051***	-0.055***	0.032***	-0.055***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Routine		-0.008***	0.248***	0.17***	-0.013***	0.088***
		(0.000)	(0.003)	(0.003)	(0.000)	(0.003)
R-Squared		0.754	0.869	0.676	0.752	0.728
F-Statistic		31,948	94,217	14,113	27,354	13,990
				1st Stage: Routine		
Abstract		-0.011***	-0.019***	-0.018***	-0.009***	-0.024***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Routine		0.067***	0.175***	0.262***	0.070***	0.196***
		(0.000)	(0.003)	(0.003)	(0.000)	(0.003)
R-Squared		0.790	0.961	0.780	0.787	0.792
F-Statistic		33,086	95,861	5,347	20,410	6,284
Fixed-Effects	Time	Yes	Yes	Yes	Yes	Yes
	Individual	Yes	Yes			
	Occupation		Yes			
	Occupation-Spell			Yes		
	Job-Spell				Yes	
	Occupation-Job-Spell					Yes
Effective Sample		2,005,966	2,005,966	2,005,966	2,005,966	2,005,966

Note 1: The coefficients are presented along with their level of significant. A coefficient concatenated with * represents a p-value ≤ .1, ** represents a p-value ≤ .05, and *** represents a p-value ≤ .01 significance. Note 2: The results are presented with robust standard errors clustered on the requisite fixed-effect level.