August 2013

Three Essays in Labor Economics

Benjamin Jack Van Kammen
University of Wisconsin-Milwaukee

THREE ESSAYS IN LABOR ECONOMICS

by

Benjamin Van Kammen

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Economics

at

The University of Wisconsin-Milwaukee

August 2013

ABSTRACT
THREE ESSAYS IN LABOR ECONOMICS

by

Benjamin Van Kammen

The University of Wisconsin-Milwaukee, 2013
Under the Supervision of Professors Scott J. Adams and Scott Drewianka

There are three chapters in this dissertation, each of which consists of a journal-length article. They are on the following subjects.

The first chapter uses ordinances in San Francisco and Washington, D.C. to measure the effects of mandated paid sick leave on employment and wages. Using the Quarterly Census of Employment and Wages, an employment increase is observed in San Francisco and Washington, D.C. relative to places without an ordinance. This evidence suggests that sick leave mandates correct a market failure of under-provision of benefits.

The second chapter uses a novel measure of distance based on the O*Net Content Model to show that information revealed by the spouses' occupations predicts divorce. Spouses that are closer in terms of their occupations' requisite knowledge are more likely to divorce, supporting the hypothesis that gains from specialization in a household renders a marriage more durable. Dissimilar spouses in terms of their occupations' activities are more likely to divorce, suggesting that each spouse brings an inclination toward certain activities to the marriage that reflects compatible preferences for joint consumption of household public goods.

The third chapter measures intertemporal earnings correlation across occupations in the U.S. using the Current Population Survey, 1971-2012. Then predictors of occupational earnings correlation are identified from among measures of occupational dissimilarity based on the O*Net database. Its findings consist of several surprisingly positive and U-shaped relationships between distance measures and measures of earnings correlation, as well as distance measures with negative estimated effects on earnings correlation.

TABLE OF CONTENTS

List of Figures

LIST OF TABLES

Sick Leave Mandates and Employment

1. Introduction

The most convincing argument in favor of paid sick leave mandates asserts that they correct a negative public health externality by encouraging contagiously ill workers to quarantine themselves at home. These merits are significant enough that most (at least 145) nations in the world have mandated sick leave laws (Heymann, Earle and Hayes 2007) with varying levels of generosity. The United States does not have one, presumably judging that employers would provide paid sick leave voluntarily to employees that value it. This reflects the concept of paid sick leave as a non-wage employment benefit that is transacted in an implicit market along with other job amenities. Equilibrium in this market could exhibit the optimal level of sick leave provision—in which case mandating a different level would likely harm employment.

Alternatively firms may under-provide sick leave if it has external benefits that are not internalized. It is obvious that co-workers contribute to good overall workplace health when they stay home with contagious illnesses. Firms have an interest in resolving this externality by paying sick workers to stay home; it limits the spread of viruses, infectious diseases, and productivity losses among co-workers—and by doing so, maximizes profits. A health externality does not need to be confined to co-workers at the same firm, though. Since firms have no incentive to internalize other firms' employees' spillover benefits, under-provision of paid sick leave follows if health externalities extend beyond individual firms. Policy makers in most other countries, as well as an increasing number of states and municipalities in the United States, believe that sick days are under-provided for this reason and that government mandates are necessary to resolve the problem. If that is the case, a mandate would pull more individuals into the labor force and increase employment.

This paper measures the employment effects of local paid sick leave mandates by employing a differences in differences (hereafter DD) strategy. Many researchers, e.g., Card and Krueger (1994), Klerman and Leibowitz (1997) and Ruhm (1998), Baker and Milligan (2008) have used DD to analyze similar policies. I compare the outcomes of two U.S. localities (San Francisco and Washington, D.C.) that have enacted local ordinances mandating paid sick leave to places that did not do so. My estimates indicate about a 1 to 3 percent increase in employment in San Francisco and Washington, D.C., after mandating sick leave, relative to the control group. Turning to wages, the evidence is mixed between positive and negative effects. Analysis of industries within each county, however, endorses the positive wage effect estimates (and positive employment, too) and shows that the benefits of the law primarily go to industries that already provided paid leave. Together these results indicate that mandated paid sick leave, indeed, corrects a market failure. From among the hypotheses for why it increases both employment and wages, a particular version of the canonical theory emerges that uniquely explains all four findings. This states that all firms receive a health benefit from the mandate, which covers the costs of compliance for the affected firms and unambiguously benefits unaffected firms.

Section 2 of this paper summarizes the relevant preceding literature on benefits mandates. It also espouses the theoretical effects of mandating employee benefits. Section 3 describes the data and methods that are used to test for the effects of a paid sick leave mandate. Section 4 conveys the results of the methods used for this endeavor, and

Section 5 summarizes the conclusions drawn from those results. Throughout the paper I use "paid sick days" and "paid sick leave" interchangeably. Though some international statues make a distinction between the terms based on the length of the absence, I refer to all illness-related paid time off as "leave".

Table 1: Local paid sick leave mandates.

City	Basic Rate and Maximum	Alternative Rate(s) and Maximum(s)	Relevant Dates
San Francisco	1/30, maximum 72 hours	Maximum 40 hours if firm size < 11 employees	Proposed 8/2006 Passed 11/2006 Implemented 2/2007
Washington, D.C.	1/37, maximum 56 hours	1/43 and maximum 40 hours if firm size < 100 employees 1/87 and maximum 24 hours if firm size < 25 employees	Proposed 5/2007 Passed 3/2008 Implemented 11/2008
Milwaukee	1/30, maximum 72 hours	Maximum 40 hours if firm size < 11 employees	Proposed 7/2008 Passed 7/2008 Overturned in State court 2011
Seattle	1/30, maximum 72 hours	1/40 and maximum 56 hours if firm size < 250 employees, maximum 40 hours if firm size < 50 employees	Proposed 6/2011 Passed 9/2011 Implemented 9/2012

All four places have passed a paid sick leave ordinance that applies to private employees in the city or district. Ordinances are used to establish the treatment group and the dates of treatment in my sample. "1/x" translates into 1 hour of paid sick leave earned for each x hours worked. Current information on proposals can be found at www.paidsickdays.org.

2. Background, Literature Review, and Discussion of Theory

A municipality may pass a law—through a referendum, for example—requiring private employers to provide sick leave to their employees at a specific rate (sick leave hours per hours worked) and subject to a maximum number of sick days. Four examples of such laws, and an overview of their provisions, are shown in Table 1. Additionally the

Federal government and 18 states have proposed mandated sick leave; four cities (and one state) have passed laws and only three have enforced them.

San Francisco passed the first law of this kind in 2006, and it took effect in February of 2007. Washington, D.C., introduced a law shortly after that, passed it, and implemented it in November of 2008. I consider these places as the treatment group because they are the only places to actually enforce mandates. Since that time, two cities—Milwaukee (2008) and Seattle (2011)—and one state, Connecticut (2011), have passed sick leave mandates. Seattle has recently enacted a mandate, but not enough time has elapsed to observe its effects in the QCEW. In fact none of these more recent mandates has been in place sufficiently to use its observations to identify treatment effects. Milwaukee and Seattle are useful places for selecting the control group, though. They are particularly good control observations since they proceeded nearly as far through the legislative process as the treatment places did. I discuss further how Milwaukee and Seattle are used in the sample selection section.

Presently consider what one should expect to find when looking at the data. The key theoretical forebear of this topic is Summers's (1989) "Simple Economics of Mandated Benefits", which allows for three possible effects on employment. First if employees value a mandated benefit equal to its cost, employees will effectively pay for it by accepting reduced wages at the identical level of employment. Second mandating a worthless benefit has the same effect as a tax on labor—shifting labor demand downward by the cost of the benefit without an off-setting shift in supply. A final possibility is mandating a benefit that is more valuable to employees than it costs—which increases employment.

Mitchell (1990) discusses other considerations in the same vein as Summers's, suggesting the possibility of substituting mandated benefits for other benefits. Her paper additionally contains a useful summary of previous estimation techniques for measuring wage-benefit substitutability. Mitchell also recognizes that a mandate affects the relative prices of labor in addition to the input price level that firms face. Low-skilled workers are more likely to be affected by a benefits mandate because they are less likely to receive fringe benefits voluntarily. Requiring employers to provide benefits to all employees does nothing for the employees that already receive said benefits but increases the costs of employing low-skilled workers. So substitution among inputs (away from low-skilled workers) might mitigate a mandate's disemployment effect. More on the wage-benefit substitution mechanism can be found in Woodbury (1983), Feldman (1993), Simon (2001), Olson (2002), and Marks (2011). Studies on specific groups like low-wage workers (Lee and Warren (1999) and Sherstyuk, Wachsman and Russo (2007)), and teens (Kaestner 1996) are also available, and the most common mandated benefit they examine is health insurance.

The evidence in this paper from San Francisco and Washington, D.C., is consistent with the second "Simple" scenario, however, I hypothesize that the costs of complying with the law are offset by its positive effect on employee health and productivity. As noted by Summers (178), using the example of health insurance, this may occur if employees' good health has externalities that firms have no interest in resolving. For example, good health may spill over to employees at different firms. As with a standard positive externality, paid sick leave is under-provided if its benefits transcend the buyer and seller. Then correcting the failure attracts people to the labor force—those at the margin in terms of their preferred mix of wage and benefits. If the mandated level of sick leave is chosen wisely, this explains why one should expect a positive employment effect.

Adverse selection could further explain the under-provision of sick leave. Specifically a mandate can have net positive employment effects if it corrects an information asymmetry—an idea proposed by Aghion and Hermalin (1990). Perhaps firms expect that being part of the subset of employers offering paid sick leave will attract unusually sickly applicants. But if all employers must offer sick leave by law, the sickly population diffuses among employers according to mutual benefit instead of pooling by the ones that offer leave voluntarily. In this manner, a sick leave mandate "destroys the choice" as well as the market failure.

On the empirical side of the topic, a survey of the value of sick leave can be found in Earle and Heymann (2006). One estimate of presenteeism costs comes from Goetzel, et al. (2004). Some of those authors' larger estimates suggest that on a per worker per year basis, illnesses result in \$100 to \$400 of lost productivity (407). But many conditions studied in that paper are chronic illnesses that are unlikely to improve with a sick leave allowance. Lovell's (2004) treatment is more of a cost-benefit analysis, but a precise estimate (measured in GDP or worker-days) of the public health benefit is absent. Even if the mandate harmed labor market efficiency, overall public health benefits are potentially great enough to offset those costs and make this a beneficial policy.

In terms of methodology and subject, the four papers that are most similar to this one are: Ruhm's (1998) study of parental leave mandates in European countries, Klerman and Leibowitz's (1997) study of maternity leave mandates in 12 American states between 1987 and 1993, Ziebarth and Karlsson's (2010) study of changes to sick leave mandates in Germany during 1996-97, and Petro's (2010) case study of the San Francisco sick leave ordinance.

The first two papers rely on DD as well as "differences in differences in differences" since they also exploit differences between males' and females' uses of parental leave (Ruhm) and between new mothers and mothers of older children (Klerman and Leibowitz) to measure the policies' effects. Petro does a less formal version of this, focusing on small firms and the retail and food service industries. All four papers use DD methods in which a group is treated with a change to its mandated level of leave, and that group is compared to another that is untreated.

Ziebarth and Karlsson emphasize labor costs more than employment levels and find significant decreases in the utilization of sick days under the less generous mandate. Ruhm's results indicate that mandated parental leave increases female EP ratio by 1.3 to 1.8 percentage points. Klerman and Leibowitz conclude that maternal leave statutes have a negligible effect on female employment. Petro's conclusion is that employment grew relatively rapidly in San Francisco county compared to five neighboring counties—and that the growth rate of business establishments (large and small) in San Francisco outpaced neighboring counties during the period following the leave mandate.

My application of DD relies on geographical comparisons over time, but predictions about the effect can be extended to particular industries. For example, disemployment is more likely if employees cannot pay for the mandated benefit because of a minimum wage. Marks (2011) and Simon and Kaestner (2004) analyze the effects of minimum wage increases on provision of fringe benefits, hypothesizing that they could be a buffer against disemployment; the reverse of this process could occur with mandated benefits. More directly the effect of the mandate should differ between industries that already give most employees paid leave, i.e., where the law is loosely binding, and those that don't. Three forces are at work: the cost of complying with a sick leave mandate, the value of sick leave to employees that get it as a result of the mandate, and the positive health externality. Firms in industries without ex ante sick leave provision would potentially experience all three; firms with sick leave already would experience only the externality. This predicts that industries with abundant paid sick leave prior to the mandate move up the supply curve as a result of the law and experience higher wages (productivity) and employment. Industries without much pre-mandate paid sick leave could experience employment and wages increases or decreases depending on the magnitudes of the three forces. Namely the externality and employees' valuation of paid leave tend to increase employment; the cost of compliance tends to decrease employment; the cost and employees' valuation tend to decrease wages, and the externality tends to increase wages. This paper tests these predictions as well.

National Compensation Survey (Bureau of Labor Statistics 2010) tables provide statistics on the prevalence of paid sick leave by industry, using the NAICS 2 digit classification. These differences are useful for testing whether industries experience the effect of mandated paid leave differently on the basis of pre-mandate prevalence of sick leave. Together these tests and the county-wide employment effect illuminate the effect of mandated sick leave even more clearly.

Finally the mandate's effect on county-level wages is a combination of the positive effect on industries with ex ante sick leave and the (likely non-positive) effect on industries without ex ante sick leave. My measurement of this was preceded by Markussen's (2011) study of Norwegian individuals using matched employee-physician data and instrumental variables methods. Markussen finds that spells of sick leave have a negative relationship with lagged measures of earnings. Another of his papers (2010) endeavors to uncover the optimal wage replacement rate for employees on sick leave. This is not interpreted as an estimate of sick leave's compensating wage differential, though, since Norwegian workers have social insurance that provides 52 weeks of paid leave. Instead sick leave's wage effects are interpreted as depreciated human capital and signals of productivity. As evidence that the San Francisco law improved public health, my findings are preceded by Drago and Lovell (2011) who, in employee surveys, found reductions in the prevalence of sick workers and of sick children attending school.

3. Methods and Data

A. Main Estimation Model

I use panel data to estimate the following model:

$$(1)\ Y_{it} = \alpha + \alpha_i + \alpha_t + \beta SICKMANDATE_{it} + \gamma_i t + \delta' X_{it} + \varepsilon_{it}.$$

In the model, "i" indexes counties and "t" indexes months. "Y" is the natural logarithm of employment in county "i" in month "t". In the subsequent regression measuring the wage effect, "Y" is the log average weekly real wage in county "i", month "t". I include county and time (year and calendar month) fixed effects, represented by α_i and α_t, respectively. "X" is a vector of control variables: log of population, log of government employment, minimum wage, and population growth variables. SICKMANDATE is an indicator for whether a mandate is or has been enacted. "β" is the effect of the mandate.

A less restrictive—but less transparent—version of (1) allows β to change depending on time elapsed since the law's enactment. For instance β_0 is the effect of the legislation going into effect; each quarter thereafter an additional indicator switches on (with coefficient β_j). The sum ($\sum\beta_j$) is then the effect of interest and has the interpretation: change in employment beginning from enactment until the end of the sample (see McCrary (2007, 330-331)). No mandate has been repealed so far, so the SICK indicators stay on once turned on. This sum of coefficients (including lags) estimates how the employment effect accumulates as time passes. This methodology follows McCrary (2007)—particularly for the interpretation of coefficients (also see Wolfers (2006) and Jacobson, Lalonde and Sullivan (1993)). The alternative specification used for the charts (Figure 2) allows beta to change depending on the number of quarters elapsed since the mandate took effect. This is the "dynamic" specification:

$$(2)\ Y_{it} = \alpha + \alpha_i + \alpha_t + \sum_j \beta_j SICKMANDATE_{it,j} + \gamma_i t + \delta' X_{it} + \varepsilon_{it}.$$

I present examples of both specifications in the main results (Tables 3-5).

The empirical challenge is that county fixed effects likely bias the estimate of β in a cross section. Performing the fixed effects transformation subtracts time invariant county effects, including time invariant indicators for being in the treatment group, and yields an unbiased estimate of β with panel data. An addition to the robustness of the estimates is that the residuals are clustered when computing the standard errors—a technique proposed by Arellano (1987). This addresses the effects of arbitrary correlation among observations of the same place. All the standard errors presented from estimates of employment and wage in this paper are of the cluster (by county or by industry-county where applicable) robust variety.

B. Industry Analysis

To reinforce the county estimates, I estimate a model at the industry-county level. This enables me to interact the treatment variables with measures of each industry's proportion of workers affected by the mandate. Specifically I estimate the following:

$$(3)\ Y_{ijt} = \alpha + \alpha_i + \alpha_j + \alpha_{ij} + \alpha_t + \beta_1 SICKMANDATE_{it} + \beta_2 SICKMANDATE_{it} * \Pr(Unaffected)_j + \gamma_j t + \gamma_i t + \delta' X_{it} + \varepsilon_{ijt}.$$

Y is the log of employment (subsequently of average wage) in industry "j" in county "i" in month "t". The mandate indicator is defined as before, but in this model it is interacted with industry characteristics—the "$Pr(Unaffected)_j$" variable.

By "unaffected" I refer to the likelihood that workers in that industry receive sick leave voluntarily. Operationally I use NCS statistics and Quarterly Census of Employment and Wages (Bureau of Labor Statistics) data to identify the 2-digit industries according to proportion of workers with voluntary paid sick leave. For the industry analysis, I test the coefficients on SICKMANDATE (β_1), the effect on an (hypothetical) industry with no pre-mandate (voluntary) sick leave, and on the sum of it along with the interaction term ($\beta_1 + \beta_2$), the effect on an industry with universal pre-mandate sick leave.

One should expect, if the theory about public health externalities is correct, that industries with universal sick leave ($Pr(Unaffected)_j = 1$) in the treatment counties should experience no wage-fringe trade as a result of the mandate, but they would notionally still get positive (productivity-enhancing) health externalities. Thus the theory predicts they will have unambiguously higher wages and employment post-mandate. An industry with universal non-provision of sick leave ($Pr(Unaffected)_j = 0$) would also experience the health externality (if anything, even more intensely) and the wage-fringe trade as well. Thus they are predicted to experience an ambiguous employment effect and an unambiguously negative wage effect—unless the public health externality is *so* large that it swamps both of the shifts described by Summers.

C. Data

Data on employment and wages are available from the QCEW, compiled by the Bureau of Labor Statistics. The QCEW contains the number of workers in each county, disaggregated by NAICS industry. Employment is observed on a monthly basis, so the data set consists of month-county observations. Average weekly wage (again per county) is observed each quarter and can also be tabulated by industry. In addition, the QCEW reports the number of government employees (local, state, and Federal separately) in each county. Other sources of data for the main estimates come from the U.S. Census (Annual Population Estimates) for birth and mortality rates, rates of domestic and international migration and the Department of Labor (Minimum Wage Laws in the States). Minimum wage is observed for the state in which county "i" is located and in the year of which month "t" is part. San Francisco has a (higher) local minimum wage and differs from other California counties in the sample.

As the time span for the sample, I use 2003 to 2009 inclusive. San Francisco's mandate was proposed in August 2006 and passed in November 2006, so a substantial number of periods are observed to establish a pre-treatment trend. Post-treatment observations can be differentiated from one another using lags of the treatment indicator. I generated quarterly indicators for lagged implementation to measure the cumulative effect of the mandate. There are 5 lagged indicators, covering 6 quarters including the quarter in which the mandate was enacted.

Table 2: Means and conditional means of key variables.

	All Counties	Treatment Counties	San Francisco	Washington, D.C.
Log Private Employment	9.569	13.004	13.016	12.992
Log Average Weekly Real Wage	6.216	6.999	7.029	6.969
Private Employment (Level)	55,786.040	444,620.800	449,992.600	439,249.000
Average Weekly Real Wage (Level)	599.592	1,287.625	1,328.821	1,246.429
Log Population	10.852	13.411	13.562	13.261
Log of Federal Gov't Employment	5.469	10.920	9.667	12.174
Log of State Gov't Employment	5.497	10.414	10.417	10.410
Log of Local Gov't Employment	7.268	9.524	10.657	8.392
Births per 1000 Population	13.102	12.501	11.099	13.902
Deaths per 1000 Population	9.714	8.471	7.734	9.209
Rate of International 'In' Migration	1.336	5.831	7.693	3.970
Rate of Domestic 'In' Migration	1.248	-8.355	-9.799	-6.910
Minimum Wage ($)	5.826	7.745	8.711	6.779

Private employment and wages are the primary dependent variables in this paper; the public sector employment, the minimum wage, and demographic variables on this table are used as control variables in some specifications. These are time series averages for the sample, 2003-2009 inclusive. Employment data is from the QCEW; population and growth rates are from the U.S. Census, and minimum wage data is from the Department of Labor.

Aside from the variables already mentioned, time trends have autonomous effects on employment. All of the economic variables discussed are time-variant; business cycles, national population growth, globalization, et al., affect them, and the trends can differ across locations. To capture additional variation over time that might confound the interpretation of the estimates, I include calendar-month indicators, and year fixed effects (indicators), and two varieties of time trends (alternately)—treatment-specific and county-specific. In the main regression model, α_t and γ capture these trends.

D. Selection of Samples

There are several ways of choosing a sample of counties for the estimation, and each has strengths and drawbacks. On one end of the continuum, the entire QCEW sample is available. On the other, a systematically matched control group consisting of the nearest neighbors (i.e., geographically or on propensity scores) to the treatment places presents itself.

Selecting a control group geographically has been performed in previous studies: famously by Card and Krueger (1994) on restaurants on opposite sides of the New Jersey-Pennsylvania border and by Petro (2010) to the San Francisco sick days mandate (used adjacent counties). Since the San Francisco and Washington, D.C., mandates were both local in scope, a geographical comparison to nearby counties seems unlikely to satisfy the requirements that Card's and Krueger's design does—which uses state legislation that can plausibly be called exogenous to the local labor markets.

If a control group of neighbors is selected for the present analysis, then, it ought to consist of neighbors in terms of propensity to be treated. This is where other sick leave mandates campaigns are useful; they reveal willingness to consider laws like San Francisco's and Washington's. I have used the other places that have proposed sick leave mandates: Milwaukee, Seattle, Philadelphia, New York, Portland (OR), Miami, and Orange County, FL (National Partnership for Women and Families 2013), several states, to select a control group for the sample using propensity scores as a check of the robustness of the findings, but the main results in this paper do not rely on a selected control group.

The main results in this paper use all counties in the QCEW, regression DD, and fixed effects estimation to resolve the correlation between SICKMANDATE and county fixed effects. This method relies on the treatment and controls having common time trends. I control for this with (alternately) treatment-specific and county-specific time trends ($\gamma_i t$).

4. Estimation Results

A. Main Results

An increase in employment accompanies paid sick leave mandates and has a magnitude of 1 to 4 percent. This effect is comparable and same-signed whether I use treatment-specific or county-specific time trends. Accounting for idiosyncratic time trends is clearly important; the most restrictive specifications, i.e., assuming (or limiting counties to) a common trend, yields the largest estimates. The less restrictive county time trends tend to produce more conservative estimates. These estimates change little with the inclusion of control variables such as the public sector employment, birth rate, mortality rate, in migration rates, and the minimum wage level, as Table 3 shows. The

Table 3: Main regression results.

	Effect on Log Private Sector Employment							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Mandate Implemented (β)	**0.0395*****	**0.0395*****	**0.0329*****	**0.0096*****	**0.0350*****	**0.0095*****	**0.0448*****	**0.0102**
	(.0025)	(.0025)	(.0063)	(.0007)	(.0047)	(.0006)	(.0064)	(.0079)
Implied % Change Employment	**4.03%**	**4.03%**	**3.34%**	**0.97%**	**3.56%**	**0.96%**	**4.58%**	
Log of Population	0.7088***	0.7088***	0.7088***	0.0940	0.6730***	0.2684***	0.6729***	0.2684***
	(.0512)	(.0512)	(.0512)	(.0987)	(.0479)	(.0844)	(.0479)	(.0844)
Time*1000		-0.4103***	-0.4104***		-0.2809***		-0.2809***	
		(.0429)	(.043)		(.0703)		(.0703)	
{Treatment=1}*Time*1000			0.1593		0.2626***		0.2037	
			(.1933)		(.0684)		(.1269)	
Minimum Wage					-0.0062***	-0.0022*	-0.0062***	-0.0022*
					(.0021)	(.0013)	(.0021)	(.0013)
Log Federal Gov't. Employment					0.0502***	0.0351***	0.0502***	0.0351***
					(.0107)	(.0074)	(.0107)	(.0074)
Log State Gov't. Employment					-0.0138***	-0.0238***	-0.0138***	-0.0238**
					(.0039)	(.0073)	(.0039)	(.0073)
Log Local Gov't. Employment					0.0107***	0.0205***	0.0107***	0.0205***
					(.0033)	(.0073)	(.0033)	(.0058)
Observations	262,920	262,920	262,920	259,790	262,920	259,790	262,920	259,790
Panels	3130	3130	3130	3130	3130	3130	3130	3130
Year, Calendar Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Population	Population	Population	Population	All	All	All	All
Time Trends	None	Common	Treatment & Control	County Specific	Treatment & Control	County Specific	Treatment & Control	County Specific
Dynamic Treatment Effect	No	No	No	No	No	No	Yes	Yes

The dependent variable in all specifications is private sector log of employment at the county level. The panels represent counties in the sample and have only been included if they have a full time series of 84 observations each (this amounts to excluding a handful of places in the state of Alaska from the sample). The coefficient of interest is "Mandate Implemented"—which corresponds to the treatment indicator for a sick leave mandate. All parentheses contain (county) cluster robust standard errors. "Control Variables" refers to log of population, population growth rates, log of government employment. In the specifications that use dynamic treatment effects, the linear combination (sum) of the treatment and its lags' coefficients is reported. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 4: Regression using California counties ("Case Study of San Francisco").

	Effect on Log Private Sector Employment							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Mandate Implemented (β)	**0.0672*** **	**0.0361** **	**0.0176** **	**0.0099*** **	**0.0587*** **	**0.0045*** **	**0.0587*** **	**0.0102**
	(.0106)	(.0171)	(.0069)	(.003)	(.0073)	(.0015)	(.0073)	(.0066)
Implied % Change Employment	**6.95%**	**3.67%**	**1.78%**	**1.00%**	**6.04%**	**0.46%**	**6.04%**	**1.02%**
Log of Population	1.633**	1.6351**	1.6287**	-0.8539	0.7059***	0.1674	0.7059***	0.1408
	(.7122)	(.7134)	(.7468)	(1.556)	(.1892)	(.2539)	(.1892)	(.2629)
Time*1000		-2.0236**	-1.4380		-0.7239***		-0.7239***	
		(.8348)	(.9167)		(.1434)		(.1434)	
Time*San Francisco*1000		0.7417	0.4858		-0.2003		-0.2003	
		(.6044)	(.4685)		(.1235)		(.1235)	
Minimum Wage			-0.0178*	0.0015				-0.0033
			(.0102)	(.0076)				(.0059)
EP Ratio Federal Govt			-0.0747	4.3842				-3.5780
			(.1847)	(4.5107)				(2.4292)
EP Ratio State Govt			-0.0298	-6.1876				-1.8831*
			(.0278)	(3.8749)				(.9446)
EP Ratio Local Govt			0.0228	1.9220				-0.0806
			(.0225)	(1.1576)				(.2223)
Sample	All CA Counties				CA Counties, Population > 500,000			
Observations	4,872	4,872	4,872	4,814	1,344	1,328	1,344	1,328
Panels	58	58	58	58	16	16	16	16
Year, Calendar Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Population	Population	All	All	Population	Population	Population	Population
Time Trends	None	Treatment & Control	Treatment & Control	County Specific	Treatment & Control	County Specific	Treatment & Control	County Specific
Dynamic Treatment Effect	No	No	No	No	No	No	Yes	Yes

These results are from regressions using San Francisco (earliest ordinance) as the only treatment county. Two samples of counties in California are used: all CA counties (columns 1-4), and CA counties with at least 500,000 population (columns 5-8). All standard errors (in parentheses) are clustered at the county level. "Control Variables" refers to log of population, population growth rates, log of government employment, and minimum wage as in the main regression. The emphasis is on the robustness of the estimated treatment effects to the geographically selected sample. These suggest that the sick leave mandate's effect is employment growth of between 1/2 and 2 percent. The estimates are also similar using the propensity-selected sample of 9 counties, but I do not report them here. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

estimates are all precise enough to reject a zero effect at 99% confidence, except for the dynamic specification of the treatment effect using county time trends. Even that point estimate (column 8) is consistent with the others in the set, though.

A particularly clear effect of mandated sick leave appears when I identify it from San Francisco County compared to other sample counties in California (Table 4). Some of the specifications in this set are larger than those that used the full sample, but the ones that use the least restrictive (county-specific) time trends tend toward the lower end of the scale. A negative effect can be rejected with high confidence, though, in all except one specification (again dynamic effects and county time trends). A 95% confidence interval for San Francisco's mandate's effect ranges from 1 to 8 tenths of a point, so it is possible that the effect is smaller than 1 percent. These conclusions are based on specification 4 on Table 4, using the following counties with large populations: Los Angeles, San Diego, Orange, Riverside, San Bernardino, Santa Clara, Alameda, Sacramento, Contra Costa, Fresno, Kern, Ventura, San Francisco, San Mateo, San Joaquin, Stanislaus. This set of estimates is particularly tidy because it excludes the national capital from the analysis, along with its unique connection to the Federal government.

To compare this to the Petro (2010) finding, his data indicate that San Francisco County's employment increased 4.7% during the same period and that employment in the five counties surrounding San Francisco decreased 2.5%. Petro's finding is that employment growth in San Francisco was 7.2 percentage points higher than neighboring counties, but the failure to account for both population growth and time trends limits what can be causally attributable to the sick leave policy. My estimates, which do allow for causal interpretation, indicate that San Francisco's paid sick leave mandate did lead the employment level to grow compared to the control counties, but by a smaller magnitude: between ½ and 2 percent.

My estimates using the log of county average wage as the outcome support a version of Summers's model's prediction. Some of them, using industry-specific and treatment-specific time trends, show that wages are negatively (employees pay for the benefit) impacted by mandated sick leave, which I report on Tables 5a (all counties) and 5b (California). Most of these negative estimates come from the specifications that use a time trend for the treatments and a common time trend for all other counties, and they range up to 7 or 8 percent in some cases. The estimates that add county-specific time trends are, again, more moderate and negative or (Table 5a, columns 5 and 6) positive! Table 5b shows this on columns 4 and 8, as well.

A wage increase is not, by itself, a rejection of the mandated benefits theory, but it does require an additional restriction on the basic hypothesis that a public health externality justifies the mandate: labor supply does not respond to the mandate. The next sub-section makes this more explicit, but employment and wages can both increase if the increase comes from movement up the supply curve by firms that already offered paid sick leave and firms that did not offer sick leave do not counteract it. Naturally there are other explanations for this finding, too, which I discuss in section 5. But none of the other explanations is supported by my industry analysis findings.

That paid sick leave mandates are positively related to employment—moreover that they are not negatively related—is the main finding of this paper. It is observed using many different specifications and samples. The remaining results in the paper

Table 5a: Effect of sick leave mandate on average weekly wage at county level.

	Effect on Log Private Sector Average Wage							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Mandate Implemented (β)	**0.0129*** **	**-0.0112*** **	**-0.0778*** **	**-0.0733*** **	**0.0077***	**0.0072***	**<0.0000**	**0.0027**
	(.0016)	(.0018)	(.0099)	(.017)	(.0046)	(.0041)	(.0063)	(.0055)
Implied % Change Wages	**1.30%**	**-1.11%**	**-7.49%**	**-7.06%**	**0.77%**	**0.72%**		
Time*1000	0.8002***	0.8001***	0.8001***	0.8243***				
	(.0244)	(.0244)	(.0244)	(.0424)				
{Treatment=1}*Time*1000		0.5738***	0.8902***	0.7008				
		(.0723)	(.2872)	(.4558)				
Sample	All Counties							
Observations	262,920	262,920	262,920	262,920	259,790	259,790	259,790	259,790
Panels	3130	3130	3130	3130	3130	3130	3130	3130
Year, Calendar Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	None	None	None	All	None	All	None	All
Time Trends	Common	Treatment & Control	Treatment & Control	Treatment & Control	County Specific	County Specific	County Specific	County Specific
Dynamic Treatment Effect	No	No	Yes	Yes	No	No	Yes	Yes

The dependent variable in these regressions is log of average weekly wage, and the sample is the same as the employment regressions on Table 3. All standard errors (in parentheses) are clustered at the county level. "Controls" refers to population growth rates, log of government employment, and minimum wage. This table contains the somewhat enigmatic results that prompt me to analyze the county wage effect further using the industry data. Two of the less restrictive (columns 5 and 6) specifications produce positive effects, however the standard theory predicts wage will decrease—at least for firms on which the law is binding—as in columns 2-4. * p<0.1; ** p<0.05; *** p<0.01.

Table 5b: Effect of sick leave mandate on average weekly wage at county level, California.

	Effect on Log Private Sector Average Wage							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Mandate Implemented (β)	**0.0015**	**-0.0994*** **	**-0.0056*** **	**.0167*** **	**-0.0002**	**-0.0869*** **	**-0.0046*** **	**0.0225*** **
	(.0016)	(.0081)	(.0013)	(.0041)	(.0032)	(.0072)	(.0015)	(.0055)
Implied % Change Wages		**-9.46%**	**-0.56%**	**1.69%**		**-8.32%**	**-0.46%**	**2.27%**
Time*1000	0.7420***	0.7446***			0.6180***			
	(.1004)	(.1006)			(.1035)			
{San Francisco=1}*Time*1000	.4200***	1.6700***			0.6478***			
	(.0904)	(.1071)			(.0725)			
Sample	All CA Counties				CA Counties, Population > 500,000			
Observations	4,872	4,872	4,814	4,814	1,344	1,344	1,328	1,328
Panels	58	58	58	58	16	16	16	16
Year, Calendar Month Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	None	None	None	All	None	None	None	All
Time Trends	Treatment & Control	Treatment & Control	County Specific	County Specific	Treatment & Control	Treatment & Control	County Specific	County Specific
Dynamic Treatment Effect	No	Yes	Yes	Yes	No	Yes	Yes	Yes

The county average weekly wage model estimated using California counties only. Two samples are used, largely confirming the results in Table 5a. Columns 2 and 6 produce large negative estimates comparable to column 4 in Table 5a. Both samples yield a similarly large result when dynamic treatment effects are estimated using two time trends (one for the treatment counties and one for non-treatment counties); this is regardless of whether the control variables are included or not. When I use the random trends model (columns 3, 4, 7, 8 on Table 5b), this large effect disappears in most specifications, and even changes signs. The difficulty when looking at these estimates in isolation is confidently signing the net wage change. All standard errors (in parentheses) are clustered at the county level. "Controls" refers to population growth rates, log of government employment, and minimum wage. * p<0.1; ** p<0.05; *** p<0.01.

attempt to identify the specific mechanism through which this occurs using the policy's effects on individual industries.

B. Industry Analysis

It is possible to distinguish among industries according to the fraction of their workers that receives paid sick leave voluntarily. As I alluded in sub-section 4A, looking at industries can inform whether the positive or the negative wage effect should be favored and how positive employment and wage effects can be explained in the county model. I test among four (listed in section 5) hypotheses using industry level (QCEW) data and statistics from the NCS. My results support a version of the canonical mandated benefits theory in which the mandate is individually costly to firms that don't already offer the benefit, employees do not respond by increasing labor supply, and individual costs are offset by the public health externality. I reach this conclusion from estimating the coefficients $\beta_1$and β_2 in the industry equation (3) with the expectation that β_1 represent the effect on an industry with no pre-mandate paid sick leave, and the linear combination, $\beta_1 + \beta_2$, represents the effect on an industry with universal pre-mandate paid sick leave. The dependent variables are logs of employment and wages at the 2 digit industry level. Standard errors in these estimates are clustered on industry-county; otherwise the model is the same as previously. Tables 7a (employment) and 7b (wage) compares the growth across industries within the treatment counties.

The prevalence of paid leave in a particular county is imperfectly correlated with the national statistics, so noise in these variables may account for the imprecision in some of these specifications. Columns 1 and 2, for instance, support the standard theory in

Table 6: Industry characteristics.

Industry	**Proportion Voluntary Sick Leave**	**Proportion at or Below Minimum Wage**	**Average Employees per Firm**	**Average Weekly Wage**
Agriculture, Forestry, Fishing and Hunting	0.36	0.042	6.305	0.500
Mining	0.54	0.065	5.709	2.230
Utilities	0.93	0.006	51.661	1.018
Construction	0.36	0.011	11.300	0.987
Manufacturing	0.61	0.019	22.894	1.062
Wholesale Trade	0.79	0.029	9.438	1.207
Retail Trade	0.51	0.073	13.602	0.546
Transportation and Warehousing	0.72	0.017	13.151	0.518
Information	0.89	0.039	27.599	1.351
Finance and Insurance	0.92	0.023	20.912	1.885
Real Estate and Rental and Leasing	0.80	0.047	8.946	0.925
Professional, Scientific, and Technical Services	0.83	0.018	12.302	1.480
Management of Companies and Enterprises	0.64	0.040	57.553	2.025
Administrative and Support and Waste Management and Remediation Services	0.40	0.040	26.034	0.585
Educational Services	0.75	0.042	54.626	0.829
Health Care and Social Assistance	0.78	0.033	24.117	0.832
Arts, Entertainment, and Recreation	0.64	0.103	22.805	0.781
Accommodation and Food Services	0.29	0.256	21.648	0.371
Other Services (except Public Administration)	0.53	0.088	4.977	0.663
All Industries Pooled	**0.62**	**0.065**	**13.571**	**1.000**

Statistics on voluntary leave and proportion at or below the minimum wage come from the 2010 March NCS. Consequently they are national averages that are not necessarily precise for each county in the sample.

which employees in industries bound by the law “pay for” sick leave with lower wages, and their willingness to do so exceeds the costs of providing them paid leave (evidenced by the positive employment effect). This is consistent with the negative county-level wage effect estimates. Unfortunately, despite point estimates that support this hypothesis, they are too imprecise to reject the null with high confidence.

That is not the case with columns 3 and 4, which feature a less restrictive (industry and county-specific) allowance for time trend heterogeneity. There is neither an employment nor a wage effect on the industries bound by the law, but there are positive employment (almost 2.5%) and wage effects (about 1%) on industries that already provided paid sick leave. So the estimates that are precise enough to sign the effects reveal that only the “unaffected” industries benefit from the law, i.e., positive employment and wage effects are observed for industries that already have paid sick leave and no such effects are observed for industries without it. This evidence points to a more specific version of mandated benefits theory: the law has a positive health externality, but the non-effects on industries most severely bound by the law would require that labor supply not shift out in response to the mandate.

Specifically a 2.4% employment increase and a 0.9 to 1.0% wage increase are observed for the hypothetical industry with universal pre-mandate sick leave. Conversely the effect is (statistically and practically) insignificantly different from zero on both variables for the hypothetical industry with no pre-mandate sick leave. Together with the results of the county model, this supports the positive county employment and wage effect and illuminates the reasons both effects occur. Paid sick leave mandates confer a public health benefit on all workers by inducing the reluctant workers and firms to utilize

Table 7a and 7b: Industry characteristics' effects on employment and wage growth.

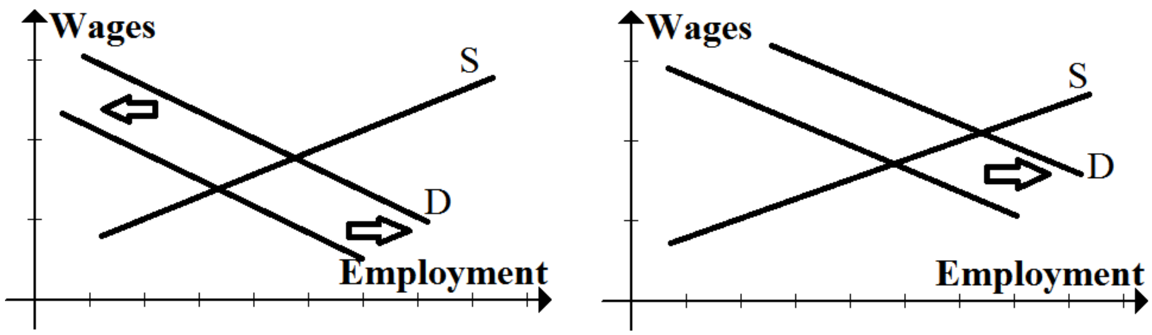
	Effect on Log Industry Employment			
	1	**2**	**3**	**4**
Mandate Implemented (β1)	**0.2912**	**0.2926**	**-0.0069**	**-0.0070**
	(.1904)	(.1903)	(.0106)	(.0106)
Implied % Change Employment	**-**	**-**	**-**	**-**
Mandate*Unaffected (β2)	-0.2136	-0.2166	0.0306**	0.0305**
	(.1779)	(.1776)	(.0142)	(.0142)
Mandate, Unaffected Industry (β1+β2)	**0.0776**	**0.0760**	**0.0237*** **	**0.0235*** **
	(.0564)	(.056)	(.006)	(.006)
Implied % Change Employment	**-**	**-**	**2.40%**	**2.37%**
Observations	2,778,552	2,778,552	2,745,474	2,745,474
Panels	33,078	33,078	33,078	33,078
Year, Calendar Month Effects	Yes	Yes	Yes	Yes
Controls	Population	All	Population	All
Time Trends	Industry, Treatment & Control	Industry, Treatment & Control	Industry, County	Industry, County
Dynamic Treatment Effect	No	No	No	No

	Effect on Log Industry Average Wage			
	1	**2**	**3**	**4**
Mandate Implemented (β1)	**-0.0345**	**-0.0327**	**0.0040**	**0.0043**
	(.0232)	(.0235)	(.0048)	(.0047)
Implied % Change Wages	**-**	**-**	**-**	**-**
Mandate*Unaffected (β2)	0.0214	0.0199	.0052	0.0051
	(.0355)	(.0359)	(.007)	(.0069)
Mandate, Unaffected Industry (β1+β2)	**-0.0130**	**-0.0129**	**.0092*** **	**.0094*** **
	(.0184)	(.0184)	(.0027)	(.0026)
Implied % Change Wages	**-**	**-**	**0.93%**	**0.94%**
Observations	2,778,549	2,778,549	2,745,470	2,745,470
Panels	33,078	33,078	33,078	33,078
Year, Calendar Month Effects	Yes	Yes	Yes	Yes
Controls	None	All	None	All
Time Trends	Industry, Treatment & Control	Industry, Treatment & Control	Industry, County	Industry, County
Dynamic Treatment Effect	No	No	No	No

Estimates of model (3) using industry data. The panels are (2 digit) industry, county combinations. Again standard errors are cluster robust around each panel (county-industry).
* $p<0.1$; ** $p<0.05$; *** $p<0.01$.

sick days. This is costly to the reluctant firms, but the costs are offset by the public health externality; exposure to fewer sick workers from other firms makes their workers healthier and more productive. Firms that already offered paid sick leave receive the benefit without incurring costs and move up the supply curve resulting in higher employment and wages.

Figure 1: Industry effects of mandated benefit with positive externality, without supply response.



The panel on the left shows the effect on an (hypothetical) industry without paid sick leave prior to the mandate. As in the canonical model, the cost of complying with the law shifts demand inward, but the costs are offset by the productivity-enhancing effects of the health externality, shifting it outward and resulting in an insignificant effect. The right panel shows the effect on an (hypothetical) industry with universal paid sick leave prior to the mandate. It experiences no costs of complying with the law but still receives the health externality, shifting demand outward and resulting in positive employment and wage effects. I hypothesize that this explains the findings in columns 5 and 6 of Table 4 (positive wage effect at county level) and in columns 3 and 4 of Tables 6a and 6b (positive employment and wage effects for industry with ex ante paid leave, no effects for industry without ex ante paid leave).

C. Robustness Checks

I have performed numerous variations of this analysis that confirms the primary result of the paper: that paid sick leave mandates increase employment. Two broad categories account for most of them: different samples and different specifications. As I mentioned in section 3D, selecting a sample on the basis of propensity to be treated is an alternative to using control variables on the sample of all counties. I have estimated the model using three different propensity score matched samples to confirm the sign and magnitude of the employment effect (see appendix for additional details). I have also

Figure 2a: Employment effect of paid sick leave mandate over time.

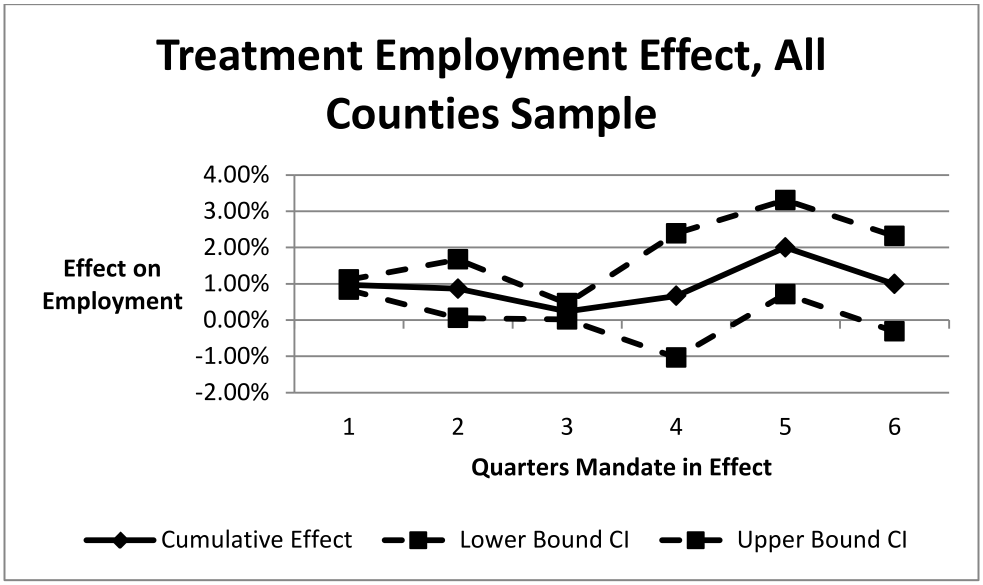
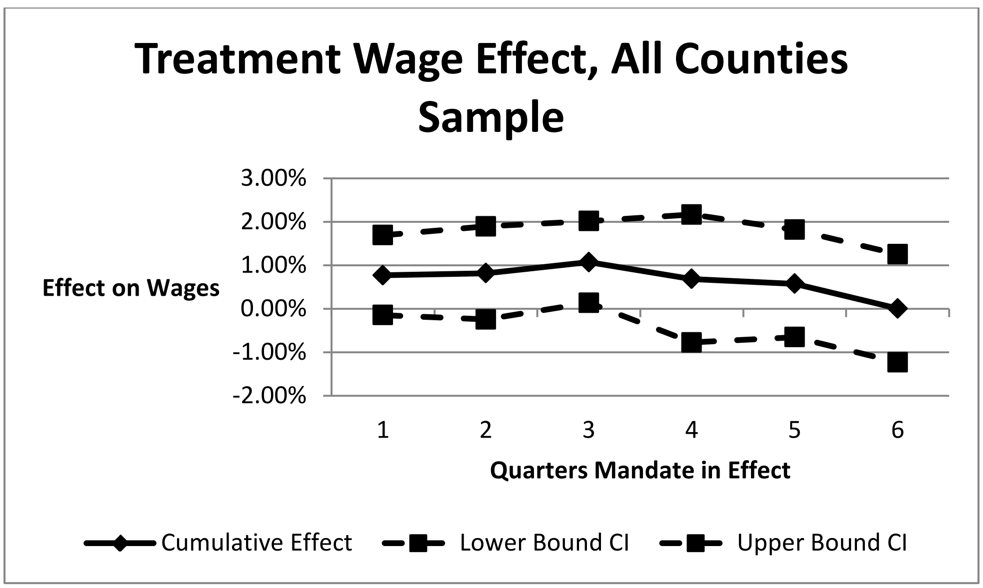


Figure 2b: Wage effect of paid sick leave mandate over time.



These figures chart the linear combinations of the coefficients on several lags of the treatment variable. The regression is specified as in the last column of Table 3 and column 7 on Table 5a, respectively. The graph also shows the boundaries of a 95% confidence interval for the effect.

estimated the model using private sector EP (employment to population) ratio as the dependent variable to confirm the primary results. I report the results that used the most transparent methods and were easiest to interpret in the paper, but the positive employment result is a feature of nearly every regression I have run using this data.

As an indication of the findings using a propensity-matched subset of counties, consider model (1) estimated using only San Francisco along with other California counties with high propensity scores as the control group: Alameda, Contra Costa, Los Angeles, Riverside, San Bernadino, Sacramento, San Diego, and Fresno. The estimated treatment effect is between (95% confidence interval) 0.4 and 0.8 EP points. San Francisco's population grew 3.6% between 2006 and 2009. It can be shown that the change in EP divided by the initial level equals: %Δ Employment minus %Δ population. Some quick calculations suggest that if mandating sick leave increases EP by one point, the level of employment increases by roughly 5.4% (initial EP ratio in San Francisco was 0.568.). Six tenths of a point of EP increase means that the level rose by roughly 4.7%.

Most of the wage effect estimates using the propensity score-matched samples are modestly negative (about 1% decrease) and follow the same pattern of the full sample: modestly positive and significant for the county-specific time trend specifications. The industry analysis also follows the same pattern, confirming that the positive county effects originate from industries with ex ante sick leave provision. Where the county wage effect estimates are negative, the industry analysis lacks precision compared to its counterpart using the whole sample. The point estimates do suggest that industries without ex ante sick leave account for the lower county-wide wages, but the coefficients are not statistically significant.

Representatives of the specifications with time-varying effects are illustrated on Figure 2. I hold them up merely as crude evidence that the employment and wage effects are concentrated primarily in the immediate wake of the mandate and gradually recede. Whether employment or wages recede to their original trends does not appear likely, but the gains right after the mandate do not continue perpetually, either. I emphasize the shape of these plots more so than levels and interpret them cautiously. They seem to suggest that the positive ("health externality") effects occur right away with the costs of compliance occurring subsequently as the cumulative effects taper off closer to zero.

I end this section with a clarification: to precisely make the "shifting supply and demand" interpretation of sick leave mandates, the units of measure for labor should be hours—not employment level. There is no measure in the QCEW of changes in hours worked. Since there is a maximum number of sick hours employees can earn, employers may have an incentive to schedule more hours per employee, crossing the threshold where they stop earning paid leave. If firms do this with their leave-mandated employees, this tends to moderate the measured effect compared to a case in which demand shifts and hours per worker is held constant. Consequently the possibility of changing hours per worker does not contradict the conclusion.

5. Discussion and Conclusions

My estimates consistently portray a mandated benefit with positive employment effects. The magnitude of this is probably between 1% and 4%, which is a noteworthy effect on a local labor market. While I don't rule out a negative wage effect—and I have estimates that employees pay for mandated sick leave with lower wages (on the order of

7 or 8 percent)—a modest positive effect of around 1% is more consistent with the industry analysis. Thus the costs of mandated sick leave are paid for either by employees or public health (productivity) gains. Either way my findings are evidence of a corrected market failure. It is understandable that paid sick leave is underprovided in the absence of a mandate. The external benefits of employees' health were not appropriately internalized prior to the mandates.

If I take the negative wage (positive employment) estimates to be the right ones, Summers's paper explains them very well, but because the same estimates reveal no effect on industries with ex ante sick leave they restrict sick leave's external benefits to the "affected" industries. This is not difficult to imagine, i.e., that workers primarily interact with and affect the health of other workers in the same broad industry, and the risk of contagion for someone from another industry on the bus or walking down the street is comparatively minor.

The simple static model strains, however, at explaining a positive effect on both outcomes. Observing positive employment and positive wage effects could be the result of one of the following hypotheses:

1) The simple model holds and the effect of the externalities are so large that they push equilibrium up the new supply curve beyond the original wage,

2) The simple model is inadequate at accounting for the preservation of human capital that results from the laws' effects on turnover, i.e., it is a static model that does not account for intertemporal effects; employees

do pay for sick leave, but gaining it keeps them around longer for their productivity and wages to increase beyond their initial levels,

3) something other than a public health externality, e.g., adverse selection, explains the findings,

4) The simple model holds, but labor supply does not respond to the laws (at least not during the time horizon examined) and the health externality offsets the costs of complying with the laws.

Discerning among these options is possible using the industry analysis. For instance if number 1 holds, the employment gains would be relatively larger for industries in which most workers do not have voluntary paid sick leave (supply and demand both increase). This is not observed. Similarly number 2 would require the effect on industries without much voluntary sick leave to be positive, as the preserved human capital pushes wages up the supply curve over time. Neither is this observed in the industry analysis. The observed positive employment and wage effect on industries with abundant voluntary paid sick leave contradicts adverse selection—resolution of which would entail a flow of workers (if anything) out of industries that already had paid sick leave. Finally, if number 4 holds, this would result in positive employment and wage effects for the industries with voluntary paid sick leave and no effect on industries without it.

The results of the industry analysis endorse this last possibility. Paid sick leave mandates increase overall employment, but the effects are concentrated in a subset of industries. Namely industries with widespread voluntary provision of sick leave account for most of the growth in employment and wages. This finding using data on industries convinces me that the positive wage effect estimates are most accurate. The former group's employment and wages do not change, whereas the latter group's increase. So to recapitulate, I find that sick leave mandates result in increased employment and wages, both originating in industries on which the law is not very binding.

Critics of sick leave mandates emphasize the mobility of capital as the source of their suspicions of negative consequences. My findings show that employers do not immediately flee a locality that enacts a mandate. A longer time span is necessary to see if they do eventually. If capital flight has not proven a severe problem at the local level, the policy would have even fewer negative consequences if implemented at a national level, at which capital is less mobile. This is especially true considering the implied magnitude of the public health benefit of more paid sick leave.

The upshot of the paper is that paid sick leave mandates do not appear to be "job-killing" regulations. I have found no significant evidence of disemployment from mandated paid sick leave. As I alluded earlier, the ultimate test of the legislation would compare labor market efficiency costs and the public health benefits. My results indicate that such a comparison would favor a mandate, since paid sick leave mandates appear to have no disemployment effects. Therefore paid sick leave mandates likely have unambiguous net social benefits, and it is increasingly apparent that the U.S. should follow other nations that mandate sick leave.

WORKS CITED

Aghion, Philippe, and Benjamin Hermalin. "Legal Restrictions on Private Contracts Can Enhance Efficiency." *Journal of Law, Economics, and Organization* 6, no. 2 (1990): 381-409.

Arellano, Manuel. "Computing Robust Standard Errors for Within-Group Estimators." *Oxford Bulletin of Economics and Statistics* 49, no. 4 (1987): 431-434.

Baker, Michael, and Kevin Milligan. "How Does Job-Protected Maternity Leave Affect Mothers' Employment?" *Journal of Labor Economics* 26, no. 4 (2008): 655-691.

Bureau of Labor Statistics. "National Compensation Survey Tables." 2010. http://bls.gov/ncs/ebs/benefits/2010/ownership/private/table21a.pdf (accessed May 24, 2011).

—. *Quarterly Census of Employment and Wages.* n.d. http://www.bls.gov/cew/#faq (accessed September 18, 2010).

Card, David, and Alan Krueger. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84, no. 4 (1994): 772-793.

Drago, Robert, and Vicky Lovell. *San Francisco's Paid Sick Leave Ordinance: Outcomes for Employers and Employees.* Institute for Women's Policy Research, 2011.

Earle, Alison, and Jody Heymann. "A Comparative Study of Paid Leave for the Health Needs of Workers and their Families Around the World." *Journal of Comparative Policy Analysis* 8, no. 3 (2006): 241-257.

Feldman, Roger. "Who Pays for Mandated Health Insurance Benefits?" *Journal of Health Economics* 11 (1993): 341-348.

Goetzel, Ron Z., and et al. "Health, Absence, Disability, and Presenteeism Cost Estimates of Certain Physical and Mental Health Conditions Affecting U.S. Employers." *Journal of Occupational and Environmental Medicine* 46, no. 4 (2004): 398-412.

Guo, Shenyang, and Mark W. Fraser. *Propensity Score Analysis.* Los Angeles: SAGE, 2010.

Heymann, Jody, Alison Earle, and Jeffrey Hayes. *The Work, Family and Equity Index.* Institute for Health and Social Policy, 2007.

Jacobson, Louis S., Robert J. Lalonde, and Daniel G. Sullivan. "Earnings Losses of Displaced Workers." *American Economic Review* 83, no. 4 (1993): 685-709.

Kaestner, Robert. "The Effect of Government Mandated Benefits on Youth." *Industrial and Labor Relations Review* 50, no. 1 (1996): 122-142.

Klerman, Jacob, and Arleen Leibowitz. "Labor Supply Effects of State Maternity Leave Legislation." In *Gender and Family Issues in the Workplace*, edited by F.D. Blau and R.E. Ehrenberg, 64-85. New York: Russell Sage Foundation, 1997.

Lee, Dwight R., and Ronald S. Jr. Warren. "Mandated Health Insurance and the Low-Wage Labor Market." *Journal of Labor Research* 20, no. 4 (1999): 505-515.

Lovell, Vicky. *No Time to be Sick: Why Everyone Suffers when Workers Don't Have Paid Sick Leave.* Institute for Women's Policy Research, 2004.

Marks, Mindy S. "Minimum Wages, Employer-Provided Health Insurance and the Non-Discrimination Law." *Industrial Relations* 50, no. 2 (2011): 241-262.

Markussen, Simen. "Health and Sick Leave: Policy Makers' Trade-offs." Working Paper, 2010.

Markussen, Simen. "The Individual Cost of Sick Leave." *Journal of Population Economics*, 2011.

McCrary, Justin. "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police." *American Economic Review* 97, no. 1 (2007): 318-353.

Mitchell, Olivia. "The Effects of Mandating Benefits Packages." *NBER Working Paper*, 1990.

National Partnership for Women and Families. *Paid Sick Days Campaigns.* 2013. htto://paidsickdays.nationalpartnership.org/site/DocServer/NP_PSD_Tracking_Doc.pdf?docIE=1922 (accessed February 5, 2013).

Olson, Craig A. "Do Workers Accept Lower Wages in Exchange for Health Benefits?" *Journal of Labor Economics* 20, no. 2 (2002): 92-114.

Petro, John. *Paid Sick Leave Does Not Harm Job Growth or Business Growth.* The Drum Major Institute, 2010.

Rosenbaum, Paul R., and Donald B. Rubin. "The Central Role of the Propensity Score in Observational Studies of Causal Effects." *Biometrika* 70, no. 1 (1983): 41-55.

Ruhm, Christopher. "The Economic Consequences of Parental Leave Mandates." *The Quarterly Journal of Economics* 113, no. 1 (1998): 285-317.

Sherstyuk, Katerina, Yoav Wachsman, and Gerard Russo. "Labor Market Effects of Employer-Provided Health Insurance." *Economic Inquiry* 45, no. 3 (2007): 538-556.

Simon, Kosali. "Displaced Workers and Employer-Provided Health Insurance: Evidence of a Wage/Fringe Benefit Trade-Off?" *International Journal of Health Care Finance and Economics*, 2001: 249-271.

Simon, Kosali, and Robert Kaestner. "Do Minimum Wages Affect Non-Wage Job Attributes?" *Industrial and Labor Relations Review* 58, no. 1 (2004): 52-70.

Summers, Lawrence. "Some Simple Economics of Mandated Benefits." *American Economic Review* 79, no. 2 (1989): 177-183.

U.S. Census Bureau. *Annual Population Estimates.* October 1, 2010. http://www.census.gov/popest/states/states.html.

U.S. Department of Labor. *Minimum Wage Laws in the States.* January 1, 2013. http://www.dol.gov/whd/minwage/america.htm.

Wolfers, Justin. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96, no. 5 (2006): 1802-1820.

Woodbury, Stephen A. "Substitution Between Wage and Nonwage Benefits." *American Economic Review* 73, no. 1 (1983): 166-182.

Ziebarth, Nicolas, and Martin Karlsson. "A Natural Experiment on Sick Pay Cuts, Sickness Absence, and Labor Costs." *Journal of Public Economics* 94 (2010): 1108-1122.

Occupational Distance and Marital Stability

> “Baby ain’t it somethin’ how we lasted this long; you and me provin’ everyone wrong; don’t think we’ll ever get our differences patched; don’t really matter cuz we’re perfectly matched.”
>
> – Paula Abdul and MC Skat Kat

1. Introduction

Previous research finds that the rise of dual-earner households is highly correlated with rising divorce rates, with ample evidence supporting causality in both directions [e.g., (Johnson and Skinner 1986), (Spitze and South 1985), (Bremmer and Kesselring 2004), (Ging and Kim 2011)]. Conditional on both spouses working, however, the determinants of a successful marriage are complex. And an expansive literature explores marital stability, as well as the related questions of spousal matching and gains from marriage. A complete review is beyond the scope of this paper, but a comprehensive review can be found in E. L. Lehrer (2003). In this paper, we use a novel approach that considers information revealed through current and past occupation choices in an empirical model of the determinants of a successful marriage, as measured by divorce. We hypothesize that the type of job one chooses, and the relative distance from the spouse’s job in terms of job content, reveals much about gains form specialization and relative preferences for household goods. This conjecture is supported by a wide range of research that has shown a correlation between underlying individual traits, both cognitive and non-cognitive, that influence occupational choice (A meta-analysis and review are contained in Sheu, et al. (2010)).

Although our paper is related to the previous literature on the relative wages and schooling of the spouses [e.g., (Lam 1988) and (Liu and Lu 2006)] and their relevance for match quality, we eschew rank-order or vertical comparisons of spouses in this paper. We instead posit a "horizontal" comparison of spousal characteristics with no necessary ordinal significance. Our variables of interest measure dissimilarity between spouses' occupations on a number of dimensions, which take the form of "distance" measures. They act like cartographic distances in that they do not convey which endpoint is at higher elevation or latitude. Consequently, these occupation measures can test theories of marital stability that are either based on similarities of preferences for household goods or dissimilarities allowing for gains from specialization. Spousal contrasts can ultimately be either good or bad for a marriage, depending upon which dimensions of the occupations the distance measures are based.

Our evidence suggests that there are two dimensions of occupation distance that affect marital stability and dissolution. First, spouses whose occupational information reveals dissimilarity in terms of knowledge are less likely to divorce or separate. The knowledge-based distance measures most likely capture what the spouses are able to produce within the household given the knowledge that each spouse's occupation requires. This supports the prediction that households that can divide tasks based on comparative advantage will be more stable. Second, spouses that are dissimilar in terms of vocational activities are more likely to divorce. The activities involved in one's chosen occupation reveal preferences for activities more generally. To the extent that couples share activities and consume household public goods, one's preferences for activities can be more or less compatible with a spouse's preferences.

We extend this analysis to single-earner households. As with the literature on wage sorting among couples, this research must overcome the obstacle of missing data—namely from spouses that are not earning a wage and have no current occupation from which to measure distance. Our approach overcomes this using longitudinal data containing information on individuals' earlier occupations. A "synthetic" distance predicted from characteristics of the spouses is also used as a robustness check. Combining information gleaned from current occupations, earlier occupations, and synthetic distances suggests that the information contained in one's choice of occupation is durable in terms of determining marital stability. Moreover, since these effects are robust across these methods, it is defensible that information contained in these occupation characteristics are orthogonal to the other factors that might influence the labor force participation-marital stability relationship.

The discussion proceeds as follows. Section 2 reviews the existing literature on marital dissolution, as well as the related literature on spousal matching. Section 3 discusses the construction of occupational distance measures. Section 4 discusses the data and methods used in the present examination of marriage and marital dissolution, and the empirical results are presented and discussed in Section 5. Section 6 concludes.

2. Background and Conceptual Framework

A. Related Literature on the Strength of Spousal Matches

The framework for analyzing formation and dissolution of marriage originates with Becker (1973, 1974), who describes the household production function with members' time and market goods as inputs. Becker speculates that the returns to scale of the production function are increasing, generating the incentives to marry, and the gains are magnified if one spouse concentrates on wage earning (providing market goods) and the other spouse concentrates on direct home production. Among the many extensions of this basic model is Weiss (1997), who relies on credit constraints to explain the same sorting mechanism: the gain from marriage comes from the spouse with lower wage-earning potential financing human capital investments for the spouse with higher wage-earning potential. The high potential spouse could not otherwise invest in this manner because of constraints on borrowing against future income. Weiss (86) shows that this gain in future income comes from specialization. Thus, two spouses with equal earning ability do not benefit from marrying one another. These models imply that the optimal pattern for pairing husbands and wives is negative sorting on wages, since it maximizes the gains from specialization (see Becker 1973, 826-828).

Sorting according to labor market productivity need not be the extent of the husband-wife matching mechanism. Mitigating spouses' earnings risks is another source of gains from marriage, but without obvious implications for how spouses sort in terms of productivity level. If each spouse faces uncertainty about the income he or she will earn, having another person in the household to insure against idiosyncratic earnings shocks makes both spouses' expected utilities higher. Naturally, the advantage of such insurance is more limited as the correlation of earnings risk between spouses grows. The implication is that dissimilar occupations or industries will more effectively mitigate the risk to household consumption from earnings instability. According to the risk-sharing theory, marriages between "dissimilarly employed" spouses generate more gains. We note that *dissimilar*, here, does not imply an ordinal ranking. We simply mean that diversification of jobs is good in terms of ensuring some income for the household.

There are several other reasons to suspect that dissimilar spousal occupations affect match strength. Lich-Tyler (2003) shows how assortative matching is based on similar preferences for household public goods in the absence of differences in skills, wages, et al. Weiss and Willis (1997) found the same basic notion holds with respect to education. Specifically the marginal effect of an interaction term between husband's and wife's education decreases the probability of divorce. The authors interpret this as the result of preference complementarity and shared consumption (316). From this literature, we postulate that preferences for goods correlate with individuals' choices of occupation. Non-wage amenities and disamenities attract individuals to occupations based on their valuations of the amenities. It is reasonable to expect, for example, two people who chose to work an outdoor job to also enjoy outdoor leisure activities.

Regardless of whether preferences and labor market traits are correlated, marriage generates gains for the spouses if non-rival household goods are consumed since any amount of the good consumed by one partner gives the other partner utility also. Lam (1988) elaborates on this possibility by exploring two consequences of a household public good—correlation in preferences for the public good and home production of the public good. The first suggests positive ("outdoor work and outdoor leisure") sorting of spouses, and the second suggests negative (specialization-based) sorting. The latter depends on the public good's production function and how complementary the spouses' time inputs are in terms of allowing for specialization.

The novel approach of this paper is to use information on occupations to learn more about spousal compatibility. Given that there is ample evidence from the previous literature that both similarities and dissimilarities draw spouses together and make marriages work, we suspect the richness of information about one's occupation can shed light on the role of dissimilarities. The risk sharing explanations for match quality and the idea that dissimilar spouses could more effectively divide tasks in the household imply that spouses with proximal occupations reap fewer gains from marriage. Alternatively spouses with occupations that are dissimilar could have a disadvantage in match quality if the gains from marriage come from preference compatibility, complementarity in household public good production, or spillover of human capital within the household. The last idea, suggested by Benham (1974), states that one spouse's earnings are enhanced by the knowledge of the other spouse, assuming that the other spouse has relevant knowledge. This would be relevant when both spouses' occupations are complements in market goods production, e.g., physician and nurse. A marriage involving two such occupations could be expected to make both spouses better at doing their individual jobs, thus generating larger gains from marriage.

B. Additional Factors Explaining Marital Dissolution

As with most papers in the literature, our aim is to measure determinants of match quality but must use divorce or separation as a proxy. This relies on the assumption that poorer match quality renders divorce more likely. Spouses gain information during the marriage about its quality and the availability of better matches (there may also be some "on the job search"). Since dissolution is costly, minor adverse realizations do not compel well-matched spouses to divorce; only marginally-well-matched couples do. The question we ask in this paper then is: "are couples with more distant occupations more likely or less likely to be marginally-well-matched (ceteris paribus)?" Since we use divorce and separation to reveal marriages that are relatively poor matches, we appeal to the existing literature on marital dissolution to identify other factors that are important to include in the analysis as controls. Weiss and Willis (1997) find that shocks to the earnings of one spouse affect the probability of divorce. Such increases to the husband's earnings stabilize the marriage while positive shocks to the wife's earnings destabilize it. We interpret this result as suggesting that an increase to the higher-earning spouse stabilizes the marriage, but an increase to the lower-earning spouse destabilizes the marriage. Kalmijn, Loeve and Manting (2007) show that among Dutch couples, the stabilizing effect of income growth for the higher earner is confirmed—but only when the higher earner is male. In households in which the dominant earner is female, growth of the wife's relative income has a destabilizing effect on the marriage. Similarity in the levels of the spouses' schooling at the time of marriage also stabilizes the match, as do higher age at marriage, duration, children, and marital assets like property. Investments in human capital after marriage have mixed consequences for marital stability. They increase the earning potential of the household but do so at the expense of household production. Moreover, the additional human capital stock provides the spouse that invests more attractive outside options (Johnson and Skinner 1986).

Demographic variables indicating the spouses' religious (Charles and Stephens 2004) homogamy significantly predict a lower probability of divorce. In some samples (Bramlett and Mosher 2002), ethnic homogamy, the wife being older than the husband, and successful parental marriages do likewise. Pre-marital cohabitation and previous marriages are positively correlated with divorce probabilities in the Bramlett and Mosher CDC report as well as in other samples (Weiss and Willis 1997, 313-15). Living in an area with high male unemployment, a greater proportion in poverty, a higher proportion receiving welfare, and lower median income each predicts higher divorce probability, according to the same CDC study, as does the race of the wife. Lehrer (2008) uses the same data (the National Survey of Family Growth) set to confirm that age-at-marriage is positively related to stability.

Job displacement, particularly layoffs (as opposed to plant closures), adversely affects marital stability (Charles and Stephens 2004). This finding has been confirmed for unemployment of husbands using Danish data (Jensen and Smith 1990). Geographical movement also tends to destabilize marriages because the motive is usually a new job for one spouse and this tends to benefit that spouse more than the other (Boyle, et al. 2008). In Norwegian households, receipt of public transfers, particularly through the wife, increases the likelihood of divorce (Tjotta and Vaage 2008). Blackburn (2003) finds that this phenomenon is unrelated to the generosity of welfare programs for single mothers, i.e., welfare programs for single mothers do not incentivize divorce for women. Finally, living in an area with greater availability of other mates increases the probability of divorce (South and Lloyd 1995) as does working in an occupation with greater availability of other mates (McKinnish 2007). This literature guides our choice of covariates in regressions.

3. Measuring the Distance Between Any Two Occupations

The innovation in this paper is construction of a measure of occupational distance that can be used on pairs of spouses to test various theories of marriage. The information for measuring occupational proximity comes from the O*Net Content Model: "The O*NET database contains several hundred variables that represent descriptors of work and worker characteristics, including skill requirements" (O*Net). The *activities*, *abilities*, *knowledge* and *skills* files contain the variables we use to measure distance between occupations.[1] Version 16.0 of the database from O*Net consists of scores, from worker and occupational expert questionnaires, assessing the relevance of the various activities, abilities, knowledge, and skills to each occupation.[2]

Relevance is measured on two (ordinal) scales for each occupational dimension: *importance* (1 to 5) and *level* (0 to 7). The importance scale is accompanied by typical linear, numeric scale language, such as "not important and "extremely important". The level scale is accompanied by "anchors" that communicate what constitutes a minimal level of performance and what constitutes a sophisticated level. For example, the anchors for ability code, "1.A.2.b.2: Multi-limb Coordination" are shown below.

Level 2 Anchor: "Row a boat"

Level 4 Anchor: "Operate a forklift truck in a warehouse"

Level 6 Anchor: "Play the drum set in a jazz band"

The ordinal nature of these data poses a practical problem, and so does the existence of two scales per variable. One might worry that the average of the scores among respondents from an occupation is meaningless except in comparison to averages for that occupation on other dimensions—or to other occupations' averages on the same dimension. A couple features of the scores ameliorate this problem, however.

[1] A summary of these is located online: http://www.onetcenter.org/dl_files/ContentModel_Detailed.pdf.
[2] "An occupation expert is a person who has several years of experience and training in an occupation. He or she has the expert knowledge required to respond to questions about the skills, knowledge and activities required for work in the occupation" (https://onet.rti.org/faq_oe.cfm#Q5).

1. A dimension on which the average respondent in an occupation scores higher than another dimension can be regarded as more important (at a more sophisticated level) to the occupation.
2. An occupation in which the average respondent scores a dimension higher than the average respondent from another occupation can be regarded as more important (higher level) to the occupation with the higher average score.

Together these features—along with a ranking of each occupation on each dimension—make it possible to compare a pair of occupations according to their places in the distributions of the various O*Net dimensions. Following this premise, we construct measures of the distance between each pair of occupations based rank, as well as the raw scores. Though the results reported in the paper use the distances based on raw scores, the results are robust to using the rank-based distances as well.

The second problem we confront is the existence of two scales per variable. There are two distinct, yet consequentially similar, options for treating them: 1) treat importance scores as separate dimensions or 2) treat them as weights. The two (in the Euclidian sense) distance measures that result from these options are calculated as follow.

$$(1)\ dist^{ij} = \left[\sum_{k \in K} (A^{ik} - A^{jk})^2\right]^{\frac{1}{2}};K \text{ includes all level and importance scores.}$$

$$(2)\ dist_{importance\ weights}{}^{ij} = \left[\sum_{k \in K} W^{ik} W^{jk} \left(A^{ik} - A^{jk}\right)^2\right]^{\frac{1}{2}}$$

$$\text{where, } W^{ik} = \frac{IMP^{ik}}{\sum_{k \in K} IMP^{ik}};K \text{ includes only level scores.}$$

We prefer the second formula—which uses the relative importance scores as weights—because it distinguishes between level and importance. Instead of counting all level and importance scores equally, the weighted version counts level scores that are important to both occupations heavily and those that are unimportant (to at least one) only slightly. Only if the two occupations differ on important characteristics will they be measured as "far away" by this measure—whereas unimportant differences could result in an overstatement of the distance as measured in number 1. Consequently this paper employs the second (importance weights) calculation of distance between occupations. Once again, however, we have estimated the divorce model using non-importance-weighted distances and the estimates are robust to this.

There are four O*Net files utilized in this exercise: abilities, activities, skills, and knowledge. A distance measure can be calculated for each of the four, as well as an "overall" measure. The usefulness of this measure inheres in evaluating the proximity of any pair of occupations' skill, ability, knowledge, and activity sets. We calculate them for every pair of occupations—as defined in the 2000 Census classification scheme. Then the measures can be matched to observed pairs of occupations (one per spouse) in any household-level micro data including spousal pairs.

4. Data and Methods

A. Data

The household-level data in which we observe marriages dissolve or endure come from the 2003, 2005, and 2007 waves of the PSID (Panel Study of Income Dynamics public use dataset). Consecutive observations of each household reveal married couples and their marital status 2 years later. A binary ("remain married" equals 0) variable for

Table 8: Summary statistics of key variables.

	Pooled		One Earners		Two Earners	
Variable	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Occupations Distance 'Abilities'	-	-	-	-	0.123	0.050
Occupations Distance 'Activities'	-	-	-	-	0.196	0.071
Occupations Distance 'Skills'	-	-	-	-	0.159	0.067
Occupations Distance 'Knowledge'	-	-	-	-	0.270	0.093
|Age(Male)-Age(Female)|	3.488	3.541	3.648	3.781	3.447	3.477
Years Married (Imputed)	15.111	12.019	20.116	14.576	13.840	10.922
Age of Man When Married	29.270	8.917	29.153	10.006	29.300	8.621
Age of Woman When Married	27.082	8.241	26.913	9.073	27.125	8.018
Female is Older (=1)	0.202	0.401	0.215	0.411	0.199	0.399
Spouses same religion (=1)	0.766	0.423	0.799	0.401	0.758	0.429
Husband is white	0.715	0.451	0.725	0.447	0.713	0.453
Husband is black	0.195	0.396	0.174	0.380	0.200	0.400
Wife is white	0.726	0.446	0.733	0.443	0.724	0.447
Wife is black	0.190	0.392	0.176	0.381	0.194	0.395
Spouses are same race (=1)	0.940	0.238	0.950	0.219	0.937	0.243
Head's Parents Intact (=1)	0.784	0.412	0.793	0.406	0.782	0.413
Years Education Head	13.396	2.625	13.168	2.901	13.454	2.548
Years Education Spouse	13.452	2.452	13.133	2.555	13.533	2.420
|Education(Head)-Education(Spouse)|	1.513	1.632	1.563	1.605	1.500	1.639
Educ(HH) *Educ(Wife)	184.180	57.460	177.896	62.081	185.777	56.130
Husband earnings in 1000s	50.261	102.447	54.941	169.437	49.072	76.637
Wife earnings in 1000s	23.650	24.963	7.357	24.370	27.790	23.372
Wife earns more (=1)	0.270	0.444	0.228	0.420	0.281	0.449
City Pop. >=500k	0.140	0.347	0.132	0.338	0.142	0.349
100k<City Pop.<500k	0.253	0.435	0.236	0.425	0.257	0.437
50k<City Pop.<100k	0.114	0.317	0.110	0.314	0.114	0.318
25k<City Pop.<50k	0.130	0.337	0.130	0.336	0.131	0.337
10k<City Pop.<25k	0.171	0.377	0.199	0.400	0.164	0.370
1 Kid	0.219	0.414	0.166	0.373	0.233	0.423
2 Kids	0.239	0.426	0.230	0.421	0.241	0.428
3 Kids	0.082	0.274	0.077	0.267	0.083	0.276
4 Kids	0.024	0.152	0.044	0.206	0.018	0.134
5 or More Kids	0.007	0.086	0.012	0.107	0.006	0.080
Owns Home	0.821	0.384	0.818	0.386	0.821	0.383
Have Debt (=1)	0.572	0.495	0.470	0.500	0.598	0.491
IRA or Pvt. Annuity (=1)	0.207	0.405	0.190	0.392	0.212	0.409
Exactly 1 spouse smokes (=1)	0.182	0.386	0.195	0.397	0.179	0.383
Moved last year (=1)	0.267	0.442	0.215	0.411	0.280	0.449
Variance (Husband's Occ.) Earnings	-	-	-	-	0.749	2.240
Variance (Wife's Occ.) Earnings	-	-	-	-	0.592	1.521
Earnings Covariance (Pair)	-	-	-	-	0.054	0.777
Head Married > Once (=1)	0.254	0.435	0.275	0.447	0.248	0.432
Same Industry (=1)	-	-	-	-	0.114	0.317
Sample Size	**2549**		**517**		**2032**	

The sample summarized in this table is the 2003 cross-section of the PSID.

marital status in the future period is the dependent variable.[3] Our sample consists of couples that are married in the current period and in which at least one reports an occupation.

The PSID contains a wealth of control variables as well; nearly all of the correlates of divorce found in earlier literature are available (or can be imputed). There are 4141 observations of married dual-earning couples in the pooled sample (those married in 2003, 2005 or both) and 1427 married households in which just one spouse works. Table 8 reports summary statistics for relevant variables from the 2003 wave.

There are a few limitations that we face when constructing control variables that are noteworthy in comparison with the rest of the literature. We do not explicitly observe pre-marital cohabitation in the sample. Also the survey only asks about marital status of the parents of the head of the household—not the spouse. So we only observe whether one of the spouses has parents that remained married during childhood. Variables indicating the receipt of Temporary Assistance to Needy Families (TANF), and "other welfare" exist in the PSID, but a very small number of respondents (18) report receiving any welfare. We have not included these variables because of the trivial extent to which they vary.

B. Methods

The empirical methods and notation follow Charles and Stephens (2004, 496-97) and Weiss and Willis (1997) closely. A couple's separation hazard at a given time,

[3] Though not reported in the paper, the results are also robust to an ordered divorce variable in which couples that separate are coded as "1", those that divorce are coded as "2" and those that remain married are coded as "0".

conditional on having remained married as long as they have, depends on the gain in utility they get when married compared to dissolution (net of costs): $V_t \equiv \text{gain}, \text{year } t$.

$$(3)\ V_t = G\big(spouses' inputs(t)\big) + \beta E_t\{max[V_{t+1}, net\ utility(t+1)|\ singletons]\}$$
$$+ (\mu_i + \varepsilon_{it}) - net\ utility(t)|singletons.$$

The gain consists of the value of marriage, in three parts: present household utility (G), an expectation of future utility, a stochastic part, expressed net of the opportunity cost of marriage. Spouses remain married when the gain is positive and dissolve the marriage otherwise. In our paper, the objects of interest are variables in the match-fixed ("quality") effect, μ_i, which makes the gains larger and the match less likely to dissolve. Under Charles's and Stephens's assumptions, the separation hazard (*S*) is a linear function (g) of the duration of the marriage, the characteristics of the spouses, and the match-fixed effects:

$$(4)\ S_{it} = g[duration_{it}, spousal\ inputs_{it}, opportunity\ costs_{it}, match\ quality_i],$$

where ($\frac{\partial S_{it}}{\partial \mu_i} < 0$), and the effects can be estimated using a probit model:

$$(5)\ Pr\big(y = y_j \big| X\big) = f[\Phi(X\beta)].$$

In equation (5), *f* is the probit function and $y_j \in \{0,1\}$. The outcome variable, y, equals 1 if the couple is separated or divorced; it equals 0 if they remain married. X is the vector of explanatory variables listed in Table 8. Charles and Stephens and Weiss and Willis attempted to control for the match-fixed effects, but their main focus was on earnings shocks. We are directly interested in measures of match-specific quality in this paper, however. Those previous authors used variables such as demographic and educational homogamy to capture match-fixed effects. The present paper can be viewed as moving this literature a step forward by incorporating additional match-quality measures based on *heterogamy* in X, specifically the spouses' occupations.

The danger in relying on dissimilarity measures based on the spouses' occupations is that features of the occupations, themselves, factor into the marital value function: occupational characteristics may contribute directly to household utility (1st term in (3)), and they may also reveal the value of a spouse's outside options (last term in (3)). For this reason two sets (husband and wife) of indicators for the spouses' occupations are included in the models we estimate. Additionally the information embodied in our distance measures could be related to the variability and co-variability of occupational earnings. These should not be confused with measurements of match quality, which is how we would like to interpret the effects of the distance measures. Consequently measures of each spouse's occupation's (intertemporal real) earnings variance and the pairwise covariance are included in the model to control for any link between the O*Net distances and correlated earnings. The March CPS (1971-2012) is used to calculate the variances and covariance statistics (King, et al. 2012).[4]

We estimate β in the probit equation (5), calculate marginal effects for the distance measures from the estimates, and the estimated marginal effects show whether having dissimilar occupations is bad for marital stability. Since the data include multiple (2003 and 2005) observations of the same households, all standard errors reported are calculated based on clusters for each household. All marginal effects reported are with respect to the probability of dissolution, i.e., positive effects are destabilizing.

[4] The annual average real earnings are calculated for each occupation-year. They are expressed as natural logs, de-meaned (cross-sectionally), and then used to calculate variance (per occupation's time series) and covariance (per pair of occupations): $s^2(occupation\ i) = (T-1)^{-1}\sum_t (y_{it} - \bar{y}_i)^2$ and $s_{ij} = (T-1)^{-1}\sum_t (y_{it} - \bar{y}_i)(y_{jt} - \bar{y}_j)$.

One of the primary challenges presented by household data is that some households have only one employed spouse. Consequently the distance between the spouses' occupations is not observed. It prompts the question: "how far away from the employed spouse's occupation *would* the non-employed spouse's occupation be *if they were to work*?" To address this problem, we attempt two separate fixes:

1. use the non-working spouse's first full time (adult) occupation in place of the current occupation when calculating distances, and
2. use the non-working spouse's first occupation to identify a probabilistic current occupation and measure distance based on the expected occupation.

Resolving the issue of single earner households is crucial because there is reason to believe that single earner households would match differently. The single earners are demonstrably specializing in labor market and home production, whereas the dual earners demonstrate shared consumption or productivity, as advanced by Clark & Kanbur or Benham. It is conceivable that the former group sorts negatively (e.g., on wages) and the latter group sorts positively.

Of the 1427 single earner observations in the sample, we can estimate 1068 of their distance measures using the first full-time occupation for the non-working spouse, and we can estimate 1070 of them using the probabilistic approach (occupation with highest probability conditional on first full-time occupation). Though both methods yield similar probit results, the second is particularly attractive since it relies on revelations of the occupations that working spouses have joined and presumably non-working spouses would join. Number two is less direct in this regard, but it allows for the possibility of career progression in the interim between the first job and the present. The probabilistic occupation is the present occupation with the highest probability of selection, conditional on the individual's first full-time occupation.

Table 9: Summary statistics for and correlation measures among distance measures.

Summary Statistics					
Distance Measure	**Number Combinations**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
Abilities	126,253	0.142	0.053	0	0.417
Activities	126,253	0.222	0.071	0	0.568
Knowledge	126,253	0.180	0.069	0	0.577
Skills	126,253	0.301	0.083	0	0.663
Overall	126,253	0.444	0.122	0	1.007
Correlation Structure					
Distance Measure	Abilities	Activities	Knowledge	Skill	Overall
Abilities	1.000				
Activities	0.701	1.000			
Skill	0.803	0.761	1.000		
Knowledge	0.590	0.647	0.656	1.000	
Overall	0.812	0.877	0.880	0.895	1.000

These statistics are calculated prior to matching the distance measures to the PSID data. Hence they are not weighted to account for the prevalence of spousal pairings in occupation, i.e., the statistics treat all pairs as equally probable and attach equal weight. Our intention when reporting the measures of association is to show that each pair of measures is positively correlated and measures dissimilarity, but several pairs, such as *Activities* and *Knowledge* are far from perfectly correlated. Those two measures are capturing different dimensions of dissimilarity.

5. Results

A. Dual Earner Households

As a convenient point of departure, we present the results of the probit divorce model using the sample of married households for which we observe distance. The coefficient on "overall distance" between spouses' occupations is not significant in this model. However if the 4 constituent distance measures are included individually, two of them (activities and knowledge) have significant coefficients. We report these results on

Table 10 in columns 1 and 2. This result is comparable if the model is specified as an ordered probit (not reported).

Table 10: Divorce model: dual earner households, marginal effects of distance measures.

	Probit	
Dependent Variable: Divorce (=1)	**1**	**2**
Overall Distance	0.0328	-
	(0.0266)	-
Ability Distance	-	0.1244
	-	(0.1040)
Activities Distance	-	0.1623**
	-	(0.0749)
Skills Distance	-	0.0755
	-	(0.0895)
Knowledge Distance	-	-0.1972***
	-	(0.0617)
Household-Year Pairs	4141	4141
Includes Controls	Yes	Yes
Includes Occupation Indicators (Both Spouses)	Yes	Yes
Log Likelihood	-629.43	-622.38
Pseudo R Squared	0.1046	0.1146

* p<0.1; ** p<0.05; *** p<0.01. All standard errors are cluster robust. Marginal effects refer to the effect on the probability of dissolution (y=1). In the interest of brevity, the coefficient estimates on other covariates are relegated to a table in the appendix.

We report the marginal effects of the distance measures on Table 10. They are calculated as the effect on the probability of dissolution (separation or divorce). These show that distance between activities of the spouses' jobs is bad for the match in terms of divorce. Distance between the required knowledge is good for the match in terms of fewer divorces. In the interest of brevity, the tables only contain the binary probit results, but we have also estimated them as ordered probits, binary logits and ordered logits, and the estimates are materially the same. The next set of estimates uses the same set of X variables, with the exceptions of the "same industry" indicator and that they use the estimated occupations of non-employed spouses to measure distance for single-earner households in the sample.

B. Single Earner Households

We next proceed to analyzing the results for single-earner households. Again, the idea in looking at households where only one person works is to provide a cleaner test of the marital theories. Households have already revealed a preference for specializing in home and market work if only one works. Thus, any remaining influence of the occupational distance measures for these households suggests inherent differences among spouses that might explain the strength of marriages.

The lack of current occupational information for one spouse, however, requires us to construct the distance measures from partial information. One strategy to deal with single earner households is to predict how far away from one another their occupations would be, conditional on their other characteristics. This would amount to an out-of-sample prediction of the distances (“distance hats”) using information from the 2 earner sub-sample. Among the significant questions about the validity of such a procedure, it ignores any information contained in the working spouse’s occupation. We observe the location of that spouse’s occupation within the space—which should reveal something about the location of the other spouse’s occupation. Some occupations are in densely populated parts of the space and are close to many other occupations, whereas others are remotely located within the space.

Our preferred strategy for treating single earner households estimates the *occupation* instead of the distance; then the distance is measured from the observed occupation to the synthetic one. The PSID contains the first full-time occupation of the respondents, and one could simply use that occupation in place of the unobserved current occupation. Estimates of the divorce model using this method are in Table 11. This result is consistent with the two earner sample in terms of the signs on the marginal effects, but the estimates are less precise. The single earner sample's estimates find a stronger destabilizing effect for activities distance, and it does not find the dual earner sample's stabilizing effect of knowledge distance.

Table 11: Divorce model using first full-time occupation and probabilistic occupation for one earner households.

Dependent Variable: Divorce (=1)	**1**	**2**	**3**
Ability Distance Marginal Effect	-0.3410	-0.1388	-0.0732
	(0.2112)	(0.1818)	(0.1710)
Activities Distance Marginal Effect	0.2332	0.1721	0.2115
	(0.1052)**	(0.1101)	(0.1106)*
Skills Distance Marginal Effect	0.3506	0.0199	-0.0533
	(0.1233)***	(0.1268)	(0.1309)
Knowledge Distance Marginal Effect	-0.0041	-0.0866	-0.1659
	(0.0982)	(0.0898)	(0.0861)*
Sample	1 Earner Married	1 Earner Married	1 Earner Married
Missing Occupation	First Full Time	Probabilistic	Probabilistic
Household-Year Pairs	1068	1070	1070
Includes Controls	Yes	Yes	Yes
Includes Occupation Indicators (Both Spouses)	Yes	Yes	Yes; Interact with Indicator for Employed (=1)
Log Likelihood	-113.76	-120.68	-117.97
Pseudo R Squared	0.3457	0.2930	0.3089

* p<0.1; ** p<0.05; *** p<0.01. Cluster (household) robust standard errors for the marginal effects in parentheses. A positive marginal effect signifies an increased probability of dissolution.

The second and third columns on Table 11 allows for the likelihood that individuals travel along paths of occupations that are predictable based on their first full-time occupations. Identifying the occupation with the highest probability conditional on the first occupation lets us assign a distance measure that more closely resembles two spouses employed in the present period. Using the probabilistic occupation yields comparable estimates to the first full-time occupation; the marginal effects are similar except for the disappearance of the destabilizing effect of skills distance. Additionally the signs on the activities and knowledge effects match the signs from the two earner sample. But their precision is sensitive to whether or not the occupation indicators switch on for spouses that probabilistically (not actually) work in the occupation (column 3).

Table 12: Divorce model using pooled sample.

Dependent Variable: Divorce (=1)	**1**	**2**
Ability Distance Marginal Effect	0.0469	0.0783
	(0.0958)	(0.0933)
Activities Distance Marginal Effect	0.1606	0.1821
	(0.0649)**	(0.0655)***
Skills Distance Marginal Effect	0.1111	0.0536
	(0.0770)	(0.0793)
Knowledge Distance Marginal Effect	-0.1714	-0.1925
	(0.0526)***	(0.0530)***
Sample	1 and 2 Earners Pooled	1 and 2 Earners Pooled
Missing Occupation	First Full Time	Probabilistic
Household-Year Pairs	5211	5213
Includes Controls	Yes	Yes
Includes Occupation Indicators (Both Spouses)	Yes	Yes
Log Likelihood	-771.48	-772.60
Pseudo R Squared	0.1203	0.1159

* $p<0.1$; ** $p<0.05$; *** $p<0.01$. Cluster (household) robust standard errors for the marginal effects in parentheses. A positive marginal effect signifies an increased probability of dissolution. The marginal effects are "grand margins". The group marginal effects are shown and tested for equality on the next table.

Table 13: Wald tests of differences by household type.

Divorce Model Interaction Terms		1	2
Activities Distance	2 Earner	0.1666	0.1774
		(0.0754)**	(0.0752)**
	1 Earner	0.1376	0.2003
		(0.1092)	(0.1149)*
	Chi Squared	0.05	0.03
Knowledge Distance	2 Earner	-0.1842	-0.2043
		(0.0595)***	(0.0598)***
	1 Earner	-0.1217	-0.1470
		(0.0925)	(0.0901)
	Chi Squared	0.35	0.31
Sample		1 and 2 Earners Pooled	1 and 2 Earners Pooled
Missing Occupation		First Full Time	Probabilistic
Household-Year Pairs		5211	5213
Includes Controls		Yes	Yes
Includes Occupation Indicators (Both Spouses)		Yes	Yes
Log Likelihood		-771.48	-772.60
Pseudo R Squared		0.1203	0.1159

This table shows the marginal effects of the two significant distance measures by household type (single and dual earners). A positive marginal effect signifies an increased probability of dissolution. Both columns are derived from the estimates on Table 12, estimating marginal effects for the two groups using group-specific covariate means. The Chi Squared statistic tests the null hypothesis that both groups' marginal effects are equal. Cluster (household) robust standard errors for the marginal effects in parentheses. The emphasis is on the non-significance of the Chi Squared statistics, which leads us to not reject the null hypothesis that distance has a common effect on marital stability for both types of households.

Lastly we present the estimates on the pooled sample using these methods to address missing distances. Since this includes both types of households, we include an indicator for the single earner sub-sample. We also interact that indicator with the four distance measures and test whether the effect of distance on marital stability differs for the two groups. This test summarizes the primary conclusion as well: regardless of whether both spouses work, similarities between their vocational activities stabilize the marriage, and similarities between their vocational knowledge destabilize the marriage. We do not reject the null hypothesis that the effect is equal for both groups in the sample (Table 13). The full set of estimates is contained on a very large table in the appendix; it includes all of the explanatory variables from the models on Table 9 and Table 12 except for occupation indicators (because of their large number).

C. Discussion

The empirical results suggest that the combination of two spouses' chosen occupations—even if one of them is essentially a counterfactual—predicts whether the marriage will dissolve. Specifically more distant occupations in terms of activities destabilize a marriage, and more distant occupations in terms of knowledge stabilize a marriage. This finding is apparent among households with either one spouse or two spouses employed as well as both groups pooled together. It is consistent across methods for treating single earner households.

What do these dissimilarity measures mean for the several theories of marriage? An inference may be made by examining the descriptions of the O*Net variables in the appendix. As the broad categories suggest, activities consist of actions workers perform on their jobs, and knowledge consists of the content information needed to perform each job successfully. Occupation choices based on items on the activities list are more reflective of *preferences,* whereas choices based on the knowledge reflect *comparative advantage*. A worker with a given set of knowledge can be expected to gravitate toward an occupation that entails performing relatively pleasurable activities. Similarly a worker with a given set of preferences over activities can be expected to choose an occupation at which he possesses masterful knowledge. Applying this interpretation to spouses' occupational distance measures, spouses who perform similar activities at work are treated as having similar preferences and spouses whose jobs require similar knowledge are treated as having similar comparative advantages.

These interpretations therefore allow a way of testing the theories of marriage. The effect of increased occupational distance on the probability of marital dissolution could have several explanations. First they could reveal something about the earnings of the occupations that transcends the individuals' observed earnings, i.e., volatility, expectations, or correlation. Characteristics of one's occupational earnings profile enter the marital gains function separately from the effect of match quality. Second occupations could reveal the values of spouses' options outside of marriage. Third they could reveal match quality directly by capturing non-redundant household capital and overlapping preferences for household goods, as described in the preceding paragraph. This is the interpretation we endorse, considering that the estimates condition on the chosen occupations themselves, using indicator variables, and their earnings variance and covariance. Consequently the risk of confusing the effects of the distances with occupational earnings effects and marital opportunity costs is minimal.

The stabilizing effect of similar activities supports theories of marriage predicated upon preferences for household (especially non-rival) goods. The results suggest that spouses are better matched when their preferences for activities overlap. We caution that preferences for work activities must be representative of preferences for goods to make this conclusion truly sound. Models of marriage based on non-rival household goods fit nicely with this result since they are particularly likely to be experiential in nature, e.g., leisure activities and spending time with children. Preferences for these non-rival household activities therefore would reasonably be assumed to be related to preferences for work activities.

Models of marriage based on specialization gain support from our findings as well. Spouses with relatively distant (non-redundant) knowledge are less likely to divorce, even if both of them work instead of fully availing themselves of specialization. If anything knowledge distance is even more important to households in which both spouses work. The finding that non-redundant knowledge benefits dual earning couples, however, suggests that there is some household production that involves both spouses and increases with the diversity of the spouses' knowledge. A more subtle question (suggested by Lam) that is not answered is whether the stabilizing effect for single earners originates from home production of a non-rival household good or whether those goods are purchased a la Lam's preliminary (471-72) model. There is no clear support for a productivity-enhancing effect of spousal knowledge—at least not directly on divorce probability. If one spouse is benefiting from the knowledge of the other, the effect on marital stability must be operating indirectly through earnings—since similar knowledge means less stability in our model.

It is also interesting to note that occupational earnings covariance stabilizes a marriage. Its effect is same-signed and statistically significant in the two earner sample and the pooled sample. This could be because both occupations are trending upward (part of the reason their earnings are correlated) and the expected earnings growth increases marital gains. Other speculations are possible, as well, but at face value this evidence downplays occupational diversification as a method for combating earnings risk in a marriage.

Finally the results of this paper speak to some of the issues raised by Lich-Tyler (2003) and Clark and Kanbur (2004), respectively: the increasing importance of preference-based matching when incomes are higher and the increasing possibility of mismatch when household public goods are relatively more important. The first follows from a de-emphasis on home production in favor of purchasing household goods as incomes rise. Since the specialization motive for matching to a spouse becomes less pronounced, it becomes increasingly important to agree with one's spouse in terms of shared consumption preferences. The second comes from spousal sorting that emphasizes the distributions of tastes among the two sexes. If the distributions do not overlap sufficiently, the outer tails of the two groups get matched together in Clark and Kanbur's model, i.e., couples with opposing preferences. These heterogeneous couples are marginally matched and vulnerable to separation. A specification including an interaction between the distance measures and household income may illuminate the first question, and a version including measures of how idiosyncratic each spouse's job is may reveal the degree of "preference mismatch".

6. Conclusions

When pop singer Paula Abdul and a cartoon cat depicting the male lead performed the song, "Opposites Attract" (1990), they were right *and* wrong about marriage. An idea as old as comparative advantage dictates that opposites attract in order to reap the greatest gains from specialization. Our findings confirm that spouses with dissimilar knowledge are better matched, other things equal. However, more usually the phrase refers to opposites on more personal dimensions. In this paper those interpersonal sources of attraction are measured as dissimilarity of activities revealed through choice of occupation. In this context, opposite preferences for activities repel, other things equal.

Previous theoretical work by economists has predicted the findings in this paper—that similar preferences likely generate substantial marital gains, but specialization in disparate tasks generates marital gains as well. Taken as a whole, the results of this study empirically support each hypothesis. The reader should be cautioned that the credible interpretations of the two significant distance measures we advance are predicated on assumptions that knowledge and preferences over activities are revealed through occupation choice and that they translate into preferences and productive inputs for household goods.

WORKS CITED

Abdul, Paula. *Opposites Attract.* Comp. Oliver Leiber. 1990.

Becker, Gary. "A Theory of Marriage: Part I." *Journal of Political Economy* 81, no. 4 (1973): 813-846.

Becker, Gary. "A Theory of Marriage: Part II." *Journal of Political Economy* 82, no. 2 (1974): 11-26.

Benham, Lee. "Benefits of Women's Education Within Marriage." *Journal of Political Economy* 82, no. 2 (1974): 57-71.

Blackburn, McKinley. "The Effects of the Welfare System on Marital Dissolution." *Journal of Population Economics* 16, no. 3 (2003): 477-500.

Boyle, Paul J., Hill Kulu, Thomas Cooke, Vernon Gayle, and Clara Mulder. "Moving and Union Dissolution." *Demography* 45, no. 1 (2008): 209-222.

Bramlett, Matthew D., and William D. Mosher. *Cohabitation, Marriage, Divorce and Re-Marriage in the United States.* Vital Health Stat 23 (22), Division of Vital Statistics, Hyattsville, MD: National Center for Health Statistics, 2002.

Bremmer, D., and R. Kesselring. "Divorce and Female Labor Force Participation: Evidence from Time Series Data and Cointegration." *Atlantic Economic Journal* 32 (2004): 174-189.

Charles, Kerwin, and Melvin Jr. Stephens. "Job Displacement Disability and Divorce." *Journal of Labor Economics* 22, no. 2 (2004): 489-522.

Clark, Simon, and Ravi Kanbur. "Stable Partnerships, Matching, and Local Public Goods." *European Economic Review* 48, no. 4 (2004): 905-925.

Conley, Timothy G., and Bill Dupor. "A Spatial Analysis of Sectoral Complementarity." *Journal of Public Economics* 111, no. 2 (2003): 311-352.

Ging, Lee Chew, and Ng Pek Kim. "The Dynamics of Divorce, Income, and Female Labor Force Participation in Singapore." *Economics Bulletin* 31, no. 4 (2011): 2757-2768.

Jensen, Peter, and Nina Smith. "Unemployment and Marital Dissolution." *Journal of Population Economics* 3, no. 3 (1990): 215-229.

Johnson, William R., and Jonathan Skinner. "Labor Supply and Marital Separation." *American Economic Review* 76, no. 3 (1986): 455-469.

Kalmijn, Matthijs, Anneke Loeve, and Dorien Manting. "Income Dynamics in Couples and the Dissolution of Marriage and Cohabitation." *Demography* 44, no. 1 (2007): 159-179.

King, Miriam, et al. *Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]*. Minneapolis, MN, November 8, 2012.

Lam, David. "Marriage Markets and Assortative Mating with Household Public Goods: Theoretical Results and Empirical Implications." *Journal of Human Resources* 23, no. 3 (1988): 462-487.

Lehrer, Evelyn. "Age at Marriage and Marital Instability: Revisiting the Becker-Landes-Michael Hypothesis." *Journal of Population Economics* 21 (2008): 463-484.

Lehrer, Evelyn L. "The Economics of Divorce." In *Marriage and the Economy: Theory and Evidence from Advanced Industrial Society*, 55-74. Cambridge: Cambridge University Press, 2003.

Lich-Tyler, Stephen. "Household bargaining, the Marriage Market, and Assortative Matching Based on Preferences." *Working Paper*, 2003.

Liu, Haoming, and Jingfeng Lu. "Measuring the Degree of Assortative Mating." *Economic Letters* 92 (2006): 317-322.

McKinnish, Terra. "Sexually Integrated Workplaces and Divorce: Another Form of On the Job Search." *Journal of Human Resources* 42, no. 2 (2007): 331-352.

O*Net. *O*Net Database.* n.d. http://www.onetcenter.org/database.html (accessed May 22, 2012).

Panel Study of Income Dynamics public use dataset. Prod. Survey Research Center, University of Michigan The Institute for Social Research. Ann Arbor, MI, 2003-2007.

Shaw, Kathryn L. "Occupational Change, Employer Change and the Transferability of Skills." *Southern Economic Journal* 53, no. 3 (1987): 702-719.

Sheu, Hung-Bin, Robert W. Lent, Steven D. Brown, Matthew J. Miller, Kelly D. Hennessy, and Ryan D. Duffy. "Testing the Choice Model of Social Cognitive Career Theory Across Holland Themes: a Meta-analytic Path Analysis." *Journal of Vocational Behavior* 76, no. 2 (2010): 252-264.

South, Scott, and Kim Lloyd. "Spousal Alternatives and Marital Dissolution." *American Sociological Review* 60, no. 1 (1995): 21-35.

Spitze, G, and S. South. "Determinants of Divorce over the Marital Life Course." *American Sociological Review* 51 (1985): 583-589.

Tjotta, Sigve, and Kjell Vaage. "Public Transfers and Marital Dissolution." *Journal of Population Economics* 21 (2008): 419-437.

Weiss, Yoram, and Robert J Willis. "Match Quality, New Information, and Marital Dissolution." *Journal of Labor Economics* 15, no. 1 (1997): 293-329.

Weiss, Yoram. *The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce.* Vol. 1, chap. 3 in *The Handbook of Population and Family Economics*, edited by M.R. Rosenzweig and O. Stark, 81-123. Elsevier Science, 1997.

Occupational Distance and Pairwise Earnings Correlation

1. Introduction

Evidence that the specificity of human capital follows occupational lines has been accumulating in the forms of returns to occupational tenure (Kambourov and Manovskii, Occupational Specificity of Human Capital 2009), earnings penalties associated with "skill switching" by displaced workers (Poletaev and Robinson 2008), and the pattern of occupational mobility (Gathmann and Schönberg 2010). This confirms what is widely assumed: that occupations are a basis for differentiation among labor market, with boundaries drawn either by an official taxonomy or according to requisite human capital common to multiple occupations. Even if each occupation requires a unique set of human capital, some pairs' (of occupations) requirements overlap more than other pairs. This prompts the question of how the wages in each occupation relate to each other. Specifically, do occupations with more similar human capital requirements have earnings that more consistently move together? This paper identifies which occupations' earnings move together over time and to what degree co-movement can be explained by measures of occupational dissimilarity (distance).

There are several reasons one would wish to know about the co-movement of wages across occupations. In addition to the value of that knowledge for studies of business cycles, it would also help workers assemble a portfolio of human capital that would help them smooth economic shocks (e.g., by maintaining skills useful in occupations that covary negatively with one another), or individuals to find a spouse equipped to reduce fluctuations in household income. However surprisingly little analysis has been performed to uncover determinants of intertemporal earnings correlation across occupations. This paper fills that void by combining two sets of statistics that are both interesting in their own right: a catalog of the correlations between occupations' (log of average annual) earnings using data from the last several decades in the U.S. and the corresponding distance measures based on occupational attributes.

In the tradition of the aforementioned authors, I use measures of dissimilarity between pairs of occupations to expose predictors of occupational earnings correlation. Specifically the measures I employ are distances between each pair of occupations' O*Net (2011) measures. These measures capture how different the requisite human capital and tasks performed are between two occupations. The hypothesis is that pairs of occupations that are different, in terms of distance measures constructed from the O*Net, have less correlated earnings because they have fewer skills in common and, hence, weaker dependence between their demand shifts. I find modest support for this hypothesis. Several distance measures reveal a statistically significant relationship with the earnings correlation measures, however their overall explanatory power is weak.

The remainder of the paper is organized as follows. Section 2 provides an overview of the theoretical bases for occupational earnings correlation and the current state of this analysis heretofore. Section 3 summarizes the data and methods used in the present analysis. Section 4 summarizes the results, and Section 5 discusses their interpretation and concludes.

2. Background and Literature Review

The closest antecedent to this paper is Conley and Dupor's (2003) analysis of industry-specific productivity growth. Their research is a natural point of departure for two reasons. First it contains a simple framework for modeling how sectors' productivity growth rates co-move. Second Conley and Dupor (C&D) utilize distance measures based on each pair of industries' input vectors, i.e., the shares of input costs paid to the other sectors. Methods used in this paper are similar to those in C&D section 4.1, in which the covariance between sectoral productivity shocks is a function of the distance measures between sectors (340-42). A salient difference between this paper and theirs is the use of occupations as the unit of analysis instead of industries.

C&D examine the consequences of stochastic technological progress in multiple sectors that grow at different rates, which dates back to Lucas and Prescott (1974), was elaborated upon by Lilien (1982), and creates "sectoral shifts" in labor markets. The shocks originate either in output demand and affect derived labor demand or in the sector's production technology directly. In either case, the consequence is sector-specific demand fluctuations and wage differentials. In a frictionless labor market, reallocation by workers would then compete away the differentials, resulting in two wage movements: up with sectoral shocks and down with entry. Sectors with co-varying wages, then, would be the consequence of contemporaneous shocks and responses. In this paper I address both main sources of contemporaneous shocks to occupations' wages. I measure how different each pair of occupations' industry allocations are; this measures the degree to which they receive common derived demand shocks. And I measure how different their human capital requirements are. This measures the extent to which they have common underlying skill content—the productive inputs that their firms employ.

The values of underlying skills, then, ultimately determine wages, e.g., the popular idea (explained eloquently by Welch (1969)) that earnings are a sum of the products of the worker's skill endowments and the prices of the skills. When technology changes such that demand for a skill increases, its price changes along with the earnings of all occupations that require the skill. Thus correlation among several occupations' demand shocks, à la C&D (328-29), reflects the degree to which their skill contents overlap. Several complications ought to be pointed out, though.

The responses to sectoral wage differentials need not be a textbook supply shift. Reder (1955) identified two channels through which sectoral shifts occur: bidding up wages to attract employees and relaxation of hiring standards. Both accomplish the shift, but they have opposing implications for wages, with the latter downgrading the composition of the occupation as an alternative to raising its wages. Which channel predominates depends on the extent to which workers of different skill levels are substitutable (more substitutable implying more down-grading). This spawned a significant literature on cyclical upgrading, of which McLaughlin and Bils (2001) provide a modern example.

Helwege (1992) explores the source of friction in responses to demand shifts, attempting to explain the durability of industry wage differentials. She finds evidence that wage differentials persist because of persistent variation in human capital across industries. The alternative theory, for which she finds no evidence, is that inter-industry differentials are only eroded by young workers entering high-paying industries and accumulating the necessary training, i.e., hiring standards are relaxed in response to the shift, and wages increase after a (training) lag. This could obscure correlation in earnings as a measure of sectoral shocks if training takes longer in different sectors. On the subject of occupational choice, though, Boskin (1974) found evidence that workers do pick occupations in this fashion, i.e., in pursuit of the highest present discounted value of expected net earnings. Moreover occupational mobility work by Kambourov and Manovskii (2009, a) finds that occupation-specific human capital is a significant source of both internal wage dispersion and trans-occupational friction.

Finally sectors need not price skills uniformly. This is a consequence of the impossibility of un-bundling a worker’s skills and selling them separately to the highest bidders, demonstrated by Heckman and Scheinkman (1987). Accordingly a technology shock for a particular skill could induce a demand shift within some, but not all, of the occupations that require the skill.

Given a measure of dissimilarity for the human capital of two occupations, it is still reasonable that the demand shifts for the occupations should be related to how distinct their requisites of human capital are. This recommends applying C&D type analysis to occupational earnings correlation. For reasons outlined above, however, distance need not predict less correlation in earnings universally. Indeed some of the findings show greater distance predicting *more* correlated earnings, as well as several U-shaped relationships between distance and earnings correlation.

3. Data and Methods

A. Data

Most of the data come from two sources: the O*Net content model and the March Current Population Survey (CPS). The calculated correlation coefficients are based on yearly observations of the average real earnings in each occupation, classified by the 1990 Census taxonomy (used to compare occupations over many years in the

CPS). The sample used is 1971 to 2012 inclusive, i.e., it extends back to when the 1970 Census taxonomy for occupations was first used. Earlier classifications do not translate sufficiently well into the uniform classification scheme used by the IPUMS CPS (King, et al. 2012) database, and inclusion of earlier years results in significant swaths of missing observations. The Integrated Public use Microdata Series (IPUMS) uses a taxonomy for occupations called “OCC1990”—which is a minor revision of the 1990 Census taxonomy—and this makes occupations observed between 1971 and 2012 uniformly classifiable. There are 386 occupations with time series observations spanning these years. Thus there are 74,691 unique correlations possible: 386 “own” correlations and 74,305 “cross” correlations.

Distances in terms of occupational attributes are the hypothesized regressors that explain earnings correlation. The regressors measure dissimilarity between two occupations in terms of the level at which workers must exhibit a given skill or perform an activity. The data on occupational distance comes from the O*Net Content Model: “The O*NET database contains several hundred variables that represent descriptors of work and worker characteristics, including skill requirements” (O*Net, O*Net Database). The *activities*, *abilities*, *knowledge* and *skills* files contain the variables to measure distance between occupations, and a summary of these is available on the web site.[1] The version 16.0 database from O*Net consists of scores, from worker and

[1] http://www.onetcenter.org/dl_files/ContentModel_Detailed.pdf.

occupational experts questionnaires, assessing the relevance of the various activities, abilities, knowledge, and skills to each occupation.[2]

Relevance is measured on two (ordinal) scales for each occupational dimension: *importance* (1 to 5) and *level* (0 to 7). The importance scale is accompanied by language such as “not important and “extremely important”. The level scale is accompanied by “anchors” that communicate what constitutes a minimal level of performance and what constitutes a sophisticated level. For example, the anchors for ability code, “1.A.2.b.2: Multi-limb Coordination” are shown below.

Level 2: “Row a boat”

Level 4: “Operate a forklift truck in a warehouse”

Level 6: “Play the drum set in a jazz band”

The ordinal and subjective nature of the data poses an empirical problem: an average of the scores among respondents from an occupation is meaningless except in comparison to averages among that occupation on other dimensions—or to other occupations’ averages on the same dimension. A couple features of the scores ameliorate this problem, however.

1. A dimension that the average respondent in an occupation scores higher than another dimension can be regarded as more important (at a more sophisticated level) to the occupation.

[2] “An occupation expert is a person who has several years of experience and training in an occupation. He or she has the expert knowledge required to respond to questions about the skills, knowledge and activities required for work in the occupation” (https://onet.rti.org/faq_oe.cfm#Q5).

2. An occupation in which the average respondent scores a dimension higher than the average respondent from another occupation can be regarded as more important to the occupation with the higher average score.

Together these features, along with a ranking of each occupation on each dimension, make it possible to compare a pair of occupations according to their places in the distributions of the various O*Net dimensions. When constructing multi-dimensional measures, the importance scales can also be used as weights to emphasize only dimensions that are important to both occupations.

There are 377 occupation categories for which earnings and distances are both observed. Therefore there are 9 occupations for which correlations are observed but not distances. This is because occupational attributes for those occupations are not reported by the O*Net.[3] Given the list of these occupations and their vague definitions, the occupational measures would be so imprecise that they would be quite uninformative. Excluding them from the analysis seems appropriate and does not harm sample size much: reducing it to 70,876 (377*376) observations.

There are two demand-side reasons for wages to move together: synchronized productivity growth and synchronized output demand shocks. The O*Net measures address the former but not the latter. To overcome this, I measure how different each pair of occupations' industry allocations are, using employer survey data available from the BLS (OES Occupational Data 2010). If the shares of two occupations' employment across industries is identical, e.g., 5% of each is in Construction, 10% of each is in

[3] The 9 occupations are: "Legislators", "Professionals not elsewhere classified", "Office machine operators not elsewhere classified", "Other telecom operators", "Mechanics and repairers not elsewhere classified", "Sheet metal duct installers", "Machine operators not elsewhere classified", "Military", and "Unknown".

Transportation, they are measured zero units away from one another. Two such occupations would experience derived demand shocks, originating from output demand shocks, in tandem. To distinguish this sort of distance from occupational content distance—which more likely reflects contrasting human capital—I employ two measures of industry employment distance. One is based on the share of each industry's employment, and the other is based on the share of the occupation's employment. Their calculation follows the Euclidean formula used to calculate occupational content distances.

B. Earnings Correlation Methods

As the dependent variable, I use the correlation coefficient of the earnings for each pair of occupations, indexed by i and j. These originate from longitudinal observations of the natural logarithm of average annual real earnings (by occupation). Each pair of occupations' time series of earnings is used to calculate the correlation of their averages over time. Additionally I perform a decomposition of the correlation that enables me to measure the portion that stems from similar time trends separately from the portion stemming from correlated residuals.

The logs of average earnings are assumed to have components that are occupation-specific (α_i), year-specific (α_t), trend idiosyncratically over time, and have stochastic fluctuations around their trends (ε_{it}).

$$(1)\ w_{it} = \alpha_i + \alpha_t + \beta_i t + \varepsilon_{it},$$

$$\text{such that } \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2).$$

An occupation's time series sample mean and the cross-sectional sample mean, respectively, $\overline{w}_i$ and $\overline{w}_t$, estimate the expectations, $\mu_i \equiv E^t(w_{it})$, and $\mu_t \equiv E^i(w_{it})$, respectively. Expressing earnings as the deviation from the cross-sectional mean ($\widetilde{w}_{it} \equiv w_{it} - \overline{w}_t$) negates year-specific effects. And expressing $\widetilde{w}$ as a deviation ("$\dot{w}_{it}$") from its time series mean is the basis for the measured correlation (see appendix B).

$$(2)\ \sigma_{ij} \equiv E^t\left[\dot{w}_{it}\dot{w}_{jt}\right] = \dot{b}_i\dot{b}_j\sigma_t^2 + Cov\left(\varepsilon_{it}, \varepsilon_{jt}\right) \text{ and,}$$

$$(3)\ \rho_{ij} \equiv Corr\left(\dot{w}_{it}, \dot{w}_{jt}\right) = \frac{\sigma_{ij}}{\left[\sigma_i^2\sigma_j^2\right]^{\frac{1}{2}}}$$

$$\text{where, } \sigma_i^2 \equiv \text{occupation i's intertemporal variance} = E^t(\dot{w}_{it})^2, \text{ and}$$

$$\dot{b}_i \equiv \beta_i - E^i(\beta_i).$$

I calculate for each pair of occupations the sample estimate (r_{ij}) of ρ_{ij} and its components in (2), which enables me to estimate the determinants of each portion separately.

There is a reason to interpret un-weighted results from this exercise with caution. The data themselves are sample means, i.e., they are calculated from CPS micro data. Consequently a pair of occupations with a large representation in the CPS and a precisely measured w_{it} is weighted the same as a pair with a noisy measurement of w_{it}. Appropriate weighting of the observations in the earnings correlation model should improve the precision of its estimates. So I calculate correlation coefficients in which the observations of average earnings are weighted by the inverse of their standard errors.

Technically r is a limited dependent variable because it takes values only on the interval [-1, 1]. Therefore it is questionable whether OLS is appropriate. As a robustness check, I estimate a logistic-transformed version of equation 5 (below) but present OLS in

Table 14: Correlation matrix of multidimensional distance measures and summary statistics.

Correlation Matrix						
Multi-Dimensional Distance Measure	Abilities	Act-ivities	Skills	Know-ledge	Shares of Industry	Shares of Occ.
Abilities	1					
Activities	0.8094	1				
Skills	0.8992	0.8481	1			
Knowledge	0.7369	0.7571	0.7561	1		
Shares of Industry Employment (Across Industries)	-0.0115	-0.0362	-0.0083	-0.0891	1	
Shares of Occupation Employment (Across Industries)	0.2624	0.2429	0.2229	0.3289	-0.0174	1

Summary Statistics						
Variable	n	Mean	Std. Dev.	Skewness	Min	Max
Abilities	70,876	0.0266	0.0096	0.2011	0	0.0569
Activities	70,876	0.0291	0.0105	0.3801	0	0.0632
Skills	70,876	0.0315	0.0136	0.4347	0	0.0768
Knowledge	70,876	0.0351	0.0104	0.0881	0	0.0702
Shares of Industry Employment (Across Industries)	70,876	0.1020	0.0962	1.7315	0	0.6539
Shares of Occupation Employment (Across Industries)	70,876	0.7211	0.2838	0.0342	0	1.4093

this paper for transparency and ease of interpretation.[4] The relationship between earnings correlation and occupational distance is not materially different, but the model fits better using the transformed LHS variable.

C. Explanatory Variables: Distance Measures

The question is which measures of distance predict correlation between two occupations' earnings. I answer this question by regressing the sample correlation coefficients (r_{ij}) on the distance measures using OLS.

$$(5)\ r_{ij} = \alpha + \sum_{m=1}^{161} \gamma_m distance_{mij} + \epsilon_{ij},$$

In (5) *i* and *j* are unique occupation pairs ($i \neq j$), *m* indexes O*Net dimensions in the set of 161 distance measures. I estimate the parameters (γ_m) in (5) with earnings correlation coefficient (or either of the components in (2)) as the dependent variable. Together this set of three estimates reveals whether each occupational distance measure explains: how strongly two occupations' earnings trend together, how strongly their yearly earnings deviations from trend synch up, and how strongly earnings synch up, overall.

The explanatory variables consist of distance measures, indicating how different each pair of occupations is in terms of the O*Net occupational attributes and in terms of their (employment) distributions across industries. They are "distances" in the sense of measuring how far away from one another the occupations are in the rankings of all occupations. Following this premise, I measure the distance between the content of each pair of occupations based on how many ranks away from one another they are on the

[4] The transformation is $logistic(r_{ij}) = \ln\left(\frac{(1+r_{ij})}{(1-\min\{r_{ij},\ 0.999\})}\right)$.

O*Net scales. For example, there are 41 *activities* dimensions (with an importance and a level scale for each). In total 161 such measures are possible using the *abilities*, *activities*, *knowledge*, and *skills* files. The (square of the) distance measure on dimension *k* for occupations *i* and *j* would be:

$$(6)\ distance_{mij} \equiv \left(A_{im} - A_{jm}\right)^2,$$

where A_{im} is the level score for occupation i on dimension m.

Since interpreting 161 coefficients individually is a challenge, I also calculate four multi-dimensional distances based on each of four O*Net files: abilities, activities, skills, knowledge. For example, the distance between two occupations' knowledge vectors would be,

$$(7)\ distance_{knowledge,ij} = \left(\sum_{k=1}^{33} imp_{ik} * imp_{jk} * distance_{kij}\right)^{\frac{1}{2}}.$$

The multi-dimensional distance calculation sums over all the dimensions in one file and weights each dimension according to the relative importance in the two occupations.

$$(8)\ imp_{ik} \equiv \left(\sum_{k=1}^{33} B_{ik}\right)^{-1} B_{ik},$$

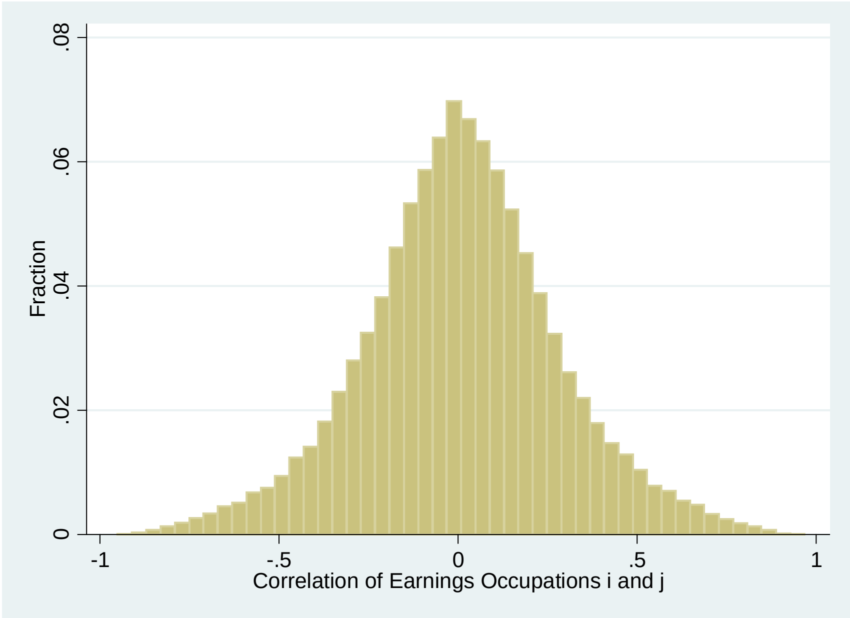
where B_{ik} is the importance score for occupation i on dimension k.

4. Results

A. Earnings Correlation Estimates

A histogram for the time series earnings correlations is shown in figure 3.

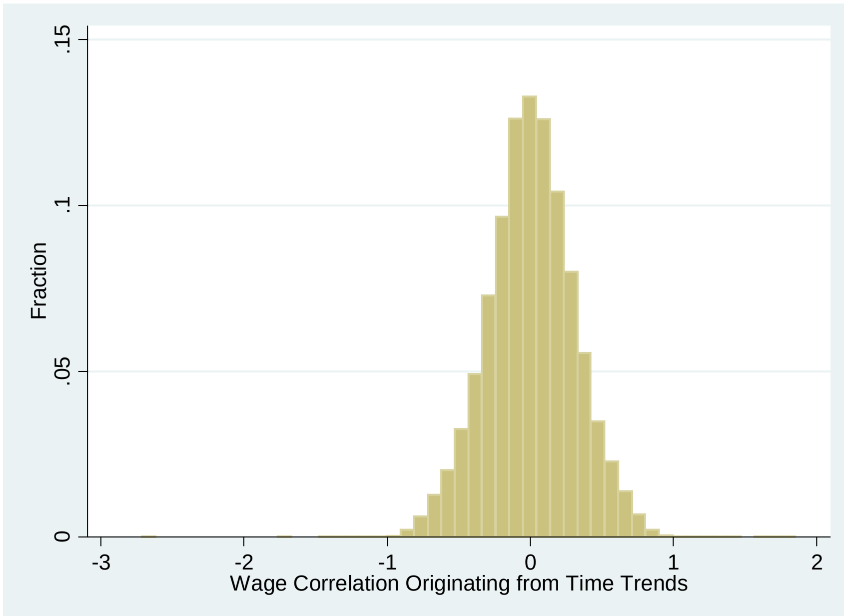
Figure 3: Distribution of the correlation coefficients, pairs of occupations' (logs of) average annual real earnings, histogram.



Given the decomposition in equation (2), the explanatory factors for the (similarity in) time trends can be estimated separately from the explanatory factors for overall earnings correlation. The distribution of the former is summarized in figure 4, and the distribution of the second component is shown in figure 5. One fact worth noting is that where the correlation coefficients are bound by the interval [-1, 1], the two components are not. Though some that fall outside the interval, such cases are rare.

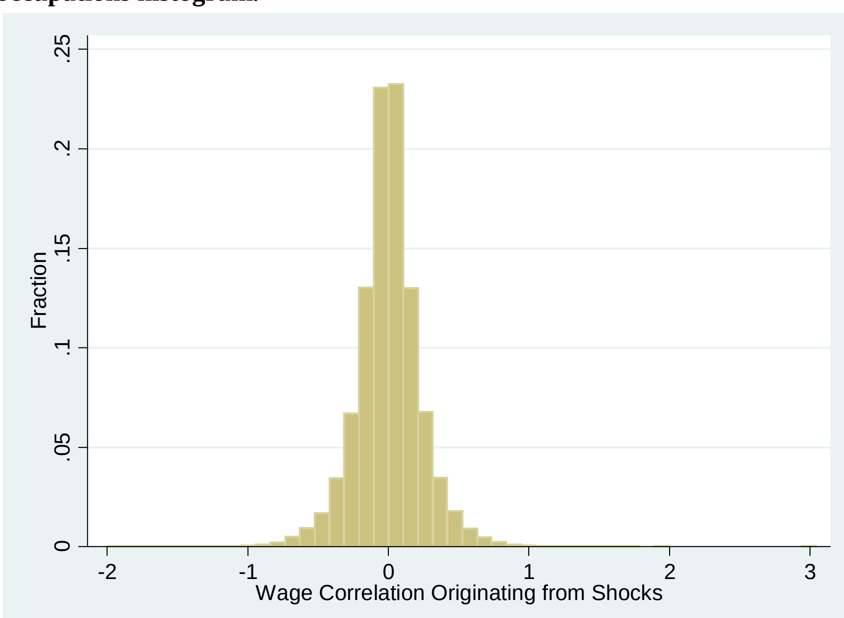
Since the number of unique correlations is large, the full set of estimates is hard to summarize concisely without narrowing the focus to a small number of occupations. Additionally the regression model in this paper is a novel attempt at making sense of this long list of correlation coefficients. The pertinent question to be answered is, "what kind of occupation pairs have correlated wages?"

Figure 4: Distribution of the time trends' components of the correlation coefficients, pairs of occupations histogram.



Note: The range of this histogram extends beyond the interval [-1,1] because this is only one of the two components in the decomposed correlation coefficient. This component, by itself, is not bound by the unit interval as long as the sum of this and the residuals component sums to [-1,1].

Figure 5: Distribution of the residuals' components of the correlation coefficients, pairs of occupations histogram.



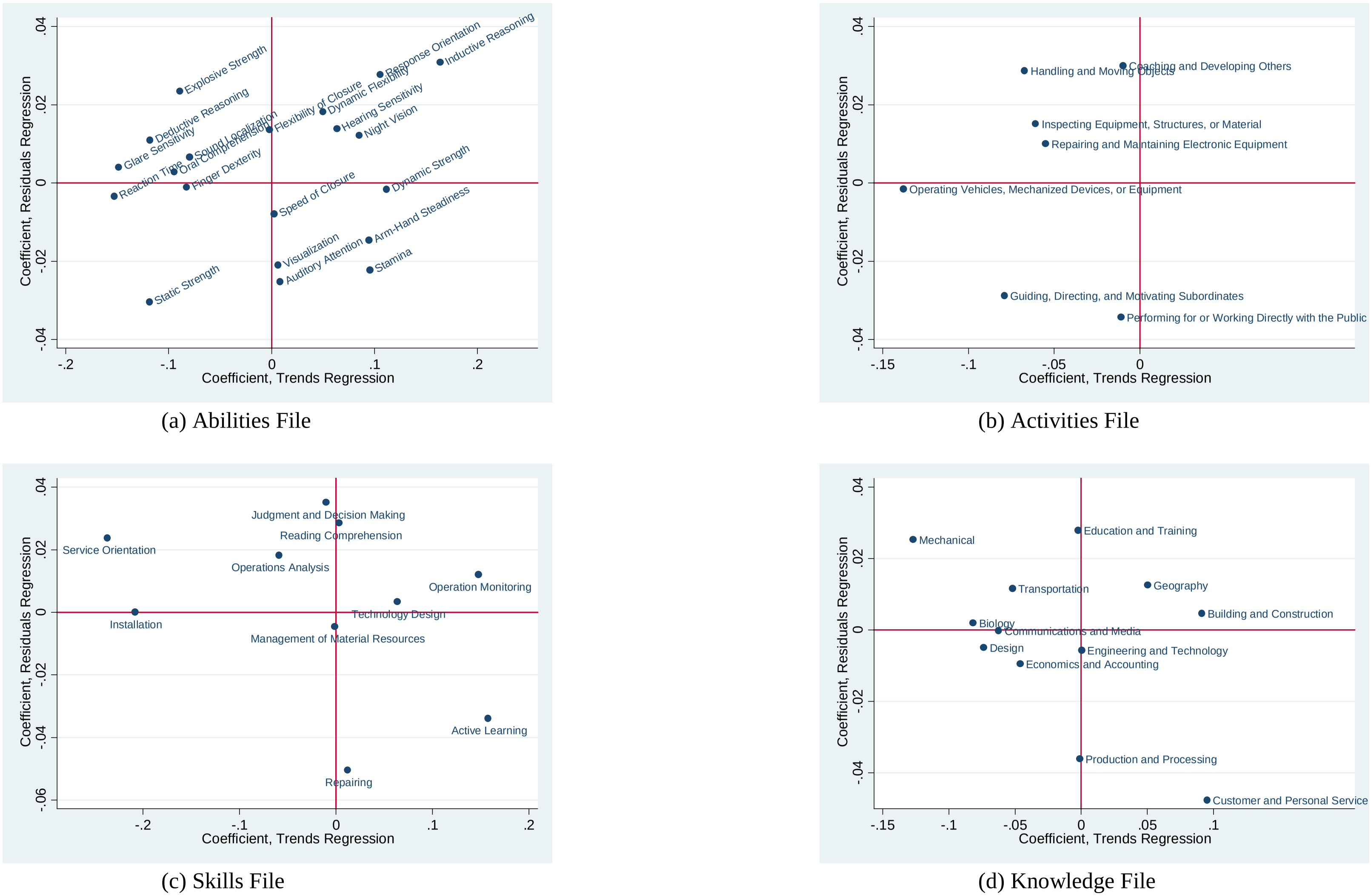
Note: The range of this histogram extends beyond the interval [-1,1] because this is only one of the two components in the decomposed correlation coefficient. This component, by itself, is not bound by the unit interval as long as the sum of this and the trends component sums to [-1,1].

B. Earnings Correlation Model

After matching the earnings correlation coefficients for occupation pairs to the corresponding distance measures, I estimate the earnings correlation model (5). The estimated coefficients and standard errors are presented on Table C1 in appendix C (first column uses the log of the transformed r_{ij}) for the earnings correlation model. Estimates are also presented in the third column for the distance measures' effects on the time trend component of correlation. And the fourth column shows the distance measures' effects on the residuals' component of correlation. A lot of the distances have coefficients that are statistically significant; this is true of all three dependent variables. For all three, the split between positive and negative is about equal. Roughly one half of the distances have coefficients that are the same-sign for both components (columns 3 and 4); among these same-signed coefficients, the split is again roughly equal between positive and negative. Despite numerous significant relationships between distance and measures of earnings correlation, the explanatory power of the model is weak, especially for the residuals component. This is revealed by low R^2 statistics in all four columns.

To graphically summarize these results, I present a scatterplot of the coefficients from the "trends" regression against the coefficients from the "residuals" regression. This illustrates which dimensions of occupational distance contribute most to earnings correlation and through which part of the decomposition they do so. The plots are divided into four groups, based on the O*Net file in which each is found. Finally the plots are restricted to include only variables with at least one *t* statistic greater than 3 in absolute value. This makes the graphs easier to read by excluding variables with imprecise coefficient estimates.

Figure 6: Scatterplots of regression coefficients from earnings correlation model, by O*Net file.



These panels plot the coefficients from the regression of the trends component against the coefficients from the regression of the residuals component. Only variables with at least one t statistic > 3 in absolute value are plotted (sub-graphs by O*Net file).

Table 15: Distance measures with largest coefficients, among one-dimensional measures.

Largest Negative Contributors to EE Correlation	Largest Positive Contributors to EE Correlation	Largest Negative Contributors to TT Correlation	Largest Positive Contributors to TT Correlation	Largest Negative Contributors to W Correlation	Largest Positive Contributors to W Correlation
Science	Judgment and Decision Making	Service Orientation	Inductive Reasoning	Service Orientation	Inductive Reasoning
Repairing	Systems Evaluation	Installation	Active Learning	Installation	Operation Monitoring
Production and Processing	Coaching and Developing Others	Reaction Time	Operation Monitoring	Reaction Time	Response Orientation
Performing for or Working Directly with the Public	Reading Comprehension	Glare Sensitivity	Dynamic Strength	Static Strength	Active Learning
Auditory Attention	Education and Training	Operating Vehicles, Mechanized Devices, or Equipment	Response Orientation	Glare Sensitivity	Dynamic Strength
		Mechanical	Stamina	Operating Vehicles, Mechanized Devices, or Equipment	Night Vision

These are the one-dimensional distance measures that have the largest coefficient estimates in the earnings correlation regression (equation 5). The list is sorted by coefficient size, but only measures with sufficiently small standard errors are included on this table. Distinctions among the four O*Net files are not made here, in terms of which file each entry comes from. This table is intended to communicate which measures have the greatest effect on earnings correlation and through which part of the decomposition they occur (residuals or trends).

The multi-dimensional distance measures allow for an easier interpretation of how dissimilarity relates to earnings correlation. They also reveal interesting non-monotonic relationships. Table 16 presents the results of regressing earnings correlation on the four multi-dimensional distances and quadratics of those distances. All four have statistically significantly non-monotonic relationships, with Abilities and Activities being the largest in magnitude. Along with Knowledge, these three have U-shaped relationships with earnings correlation. Distance between occupations initially means less correlation, but then a minimum is reached and far away occupations' earnings become *more* correlated with distance. Skills-related distance has the opposite shape (concave), reaching a maximum in the irrelevant (negative) range; therefore it is monotonically decreasing on the positive interval. But it is the weakest predictor of the four.

The industry distribution distance based on occupation employment shares exhibits a U-shaped relationship with earnings correlation, however, the minimum occurs in the negative range, so its earnings correlation is monotonically increasing in this distance (over the positive range). The analogous measure based on the shares of industry employment exhibits an inverted U-shape, and is decreasing over the positive range. This is the least surprising finding: two sets of industry shares that are different from one another means the two occupations' earnings are less correlated.

5. Discussion and Conclusions

So far this research has been exploratory in nature. I have not tested an explicit theoretical prediction of which distance measures should explain earnings correlation and why. Generally my expectation is that dissimilarity makes two occupations' earnings

Table 16: Multi-dimensional distances' relations to earnings correlation.

Multi-Dimensional Distance Measure	**1**	**2**	**3**	**4**
Abilities	-1.7685	-4.5565	-3.9326	-9.7077
	(0.2504)***	(0.9288)***	(0.5645)***	(2.0942)***
Activities	-0.4703	-3.9519	-0.9751	-8.6074
	(0.1949)**	(0.7254)***	(0.4393)**	(1.6355)***
Skills	0.4760	1.5237	1.0108	2.9788
	(0.1969)**	(0.6396)**	(0.4440)**	(1.4421)**
Knowledge	-0.7425	-2.5981	-1.6381	-6.1058
	(0.1641)***	(0.6813)***	(0.3700)***	(1.5361)***
Shares of Industry Employment (Across Industries)	-0.0366	0.0286	-0.0796	0.0692
	(0.0108)***	(0.0294)	(0.0244)***	(0.0663)
Shares of Occupation Employment (Across Industries)	-0.0355	-0.1120	-0.0800	-0.2448
	(0.0038)***	(0.0175)***	(0.0086)***	(0.0395)***
Abilities $Distance^2$		57.2940		119.9830
		(16.1768)***		(36.4733)***
Activities $Distance^2$		55.6364		122.0760
		(10.9183)***		(24.6172)***
Skills $Distance^2$		-21.0526		-41.7887
		(8.7824)**		(19.8013)**
Knowledge $Distance^2$		30.0237		71.5185
		(8.9717)***		(20.2281)***
Industry $Shares^2$		-0.1510		-0.3445
		(0.0734)**		(0.1656)**
Occupation $Shares^2$		0.0547		0.1184
		(0.0116)***		(0.0262)***
Dependent Variable	Earnings Correlation	Earnings Correlation	(Log of) Logit Transformed Correlation	(Log of) Logit Transformed Correlation
R Squared	0.01	0.01	0.01	0.01
n	70,876	70,876	70,876	70,876

* p<0.1; ** p<0.05; *** p<0.01. These are the estimates of the earnings correlation model using multi-dimensional distances. The emphasis in this table is on the shape of the relationships between distance and earnings correlation. In each case there is significant non-monotonicity. Abilities, Activities, and Knowledge each exhibit a U-shaped relationship with the earnings correlation, as do the industry allocation distances. All the O*Net distances have been expressed as a fraction of 1000.

less positively correlated, but it seems unlikely to make them more *negatively* correlated. This suggests, though, that a non-monotonic relationship may exist, and indeed I find evidence of that using multi-dimensional distances. Distances based on occupational abilities, activities, and knowledge exhibit a U-shaped relationship with earnings correlation. This finding is novel compared to C&D's finding of monotonicity among industries: ". . . covariance patterns . . . appear dictated by [input-based] distances . . . covariance declines as [input-based] distances grow." (Conley and Dupor 2003).

If occupations' labor markets mimicked C&D's (2003) spatially correlated industries, pairs of occupations would experience common demand shifts owing to productivity changes that affect the human capital general to both occupations. Then the more overlapping are their human capital requirements, the more correlation in demand shifts for the two occupations. My finding of a U-shaped relationship between earnings correlation and distance suggests that overlapping human capital requirements is not the whole story. It is tempting to conclude that the non-monotonicity reflects non-redundant and therefore complementary human capital embodied in far distant occupations. Accordingly a productivity increase for one would affect the demand for both occupations. This conclusion, however, downplays the complexity of supply responses discussed in Section 2. Especially since the explanatory power of the model is small, it is dubious that occupations experience frictionless spatially dependent sectoral shifts. Consequently I am reluctant to endorse the interpretation that the findings signal productive complementarity without qualification.

There are other reasons to interpret these findings with care. First there is employees' expectations of the intertemporal earnings profile in each occupation, i.e., climbing or declining. This idea stems from Helwege's (1992) paper, in which she reminds us that new entrants will require (pay) a premium to enter sectors with anticipated declining (climbing) future earnings. That paper is about industries, but the reasoning applies to occupations: more similarity between a pair of them suggests similar anticipated earnings streams. It is clear neither how efficient employees' expectations are nor to what extent they can act on predictable (a priori) earnings trends, but it's just one more possible source of wage differentials to obscure the effects of sectoral shifts.

With those caveats in mind, though, there are several useful lessons from the findings. I have identified occupational attributes on which dissimilarity predicts less earnings correlation. This is informative for employees who would like to diversify their human capital, e.g., if one's present résumé demonstrates only a modest degree of "Social Perceptiveness," he has an incentive to invest in this skill because occupations that require it tend to be "countercyclical" to those that do not (his present occupation).[5]

Another significant application for these results is marital stability. Risk-sharing theories of marriage (Weiss 1997) imply that household earnings risks can be reduced by diversifying, i.e., spouses choosing jobs with uncorrelated shocks. Measuring correlation between the average incomes of two spouses' occupations help identify the effect of having un-diversified earnings risks on the probability of marital dissolution.

Simply measuring the pairwise correlation between occupations' earnings is an exercise that bears fruit by itself, and several extensions are conceivable. The present paper considers the entire period (1971-2012) to estimate earnings correlation. But this

[5] Social Perceptiveness: "Being aware of others' reactions and understanding why they react as they do" (O*Net).

period could be analyzed in separate parts and used to observe changes in the degree of correlation in earnings. Interesting questions about the effects of de-unionization, female labor force participation, and international trade liberalization could be answered by examining earnings correlations based on subsamples of the CPS, e.g., before and after enactment of NAFTA.

WORKS CITED

Boskin, Michael J. "A Conditional Logit Model of Occupational Choice." *Journal of Political Economy* 82, no. 2 (1974): 389-398.

Bureau of Labor Statistics. *OES Occupational Data.* 2010. http://www.bls.gov/emp/ep_data_occupational_data.htm (accessed April 11, 2013).

Conley, Timothy G., and Bill Dupor. "A Spatial Analysis of Sectoral Complementarity." *Journal of Political Economy* 111, no. 2 (2003): 311-352.

Gathmann, Christina, and Uta Schönberg. "How General is Human Capital? A Task-Based Approach." *Journal of Labor Economics* 28, no. 1 (2010): 1-49.

Heckman, James, and Jose Scheinkman. "The Importance of Bundling in a Gorman-Lancaster Model of Earnings." *The Review of Economic Studies* 54, no. 2 (1987): 243-255.

Helwege, Jean. "Sectoral Shifts and Interindustry Wage Differentials." *Journal of Labor Economics* 10, no. 1 (1992): 55-84.

Kambourov, Gueorgui, and Iourii Manovskii. "Occupational Mobility and Wage Inequality." *The Review of Economic Studies* 76, no. 2 (2009): 731-759.

Kambourov, Gueorgui, and Iourii Manovskii. "Occupational Specificity of Human Capital." *International Economic Review* 50, no. 1 (2009): 63-115.

King, Miriam, et al. *Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]*. Minneapolis, MN, November 8, 2012.

Lilien, David M. "Sectoral Shifts and Cyclical Unemployment." *Journal of Political Economy* 90, no. 4 (1982): 777-793.

Lucas, Robert E. Jr., and Edward C. Prescott. "Equilibrium Search and Unemployment." *Journal of Economic Theory* 7, no. 2 (1974): 188-209.

Maddala, G.S. *Introduction to Econometrics.* Third. Chichester: Wiley, 2001.

McLauglin, Kenneth J., and Mark Bils. "Interindustry Mobility and the Cyclical Upgrading of Labor." *Journal of Labor Economics* 19, no. 1 (2001): 94-135.

O*Net. *O*Net Database.* July 2011. http://www.onetcenter.org/database.html (accessed May 22, 2012).

—. *O*Net Database.* n.d. http://www.onetcenter.org/database.html (accessed May 22, 2012).

Poletaev, Maxim, and Christopher Robinson. "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000." *Journal of Labor Economics* 26, no. 3 (2008): 387-420.

Reder, M.W. "The Theory of Occupational Wage Differentials." *American Economic Review* 45, no. 5 (1955): 833-852.

Weiss, Yoram. "The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce." Chap. 3 in *Handbook of Population and Family Economics*, edited by M.R. Rosezweig and O. Stark, 82-120. Elsevier Science, 1997.

Welch, Finis. "Linear Synthesis of Skill Distribution." *Journal of Human Resources* 4, no. 3 (1969): 311-327.

Wooldridge, Jeffrey M. "Models with Individual Slopes." In *Econometric Analysis of Cross-Section and Panel Data*, 315-317. Cambridge, MA: MIT Press, 2002.

Appendix A: Divorce Model Full Set of Estimates

Appendix Table A1: Coefficient estimates on control variables.

	Probit		Probit	
Dep. Variable: Divorce (=1)	**1**	**2**	**3**	**4**
Age Gap Absolute Value	0.0022	0.0023	0.0025	0.0027
	(0.0012)*	(0.0012)**	(0.0010)***	(0.0009)***
Years Married (Imputed)	(0.0016)	(0.0016)	(0.0017)	(0.0016)
	(0.0004)***	(0.0004)***	(0.0003)***	(0.0003)***
Age of Man When Married	(0.0010)	(0.0011)	(0.0015)	(0.0015)
	(0.0011)	(0.0011)	(0.0009)*	(0.0009)*
Age of Woman When Married	0.0001	0.0001	0.0005	0.0005
	(0.0011)	(0.0011)	(0.0009)	(0.0009)
Female Older	(0.0050)	(0.0053)	(0.0017)	(0.0006)
	(0.0109)	(0.0108)	(0.0092)	(0.0092)
Spouses are same religion	(0.0005)	(0.0004)	(0.0040)	(0.0041)
	(0.0066)	(0.0066)	(0.0060)	(0.0059)
Husband is white	(0.0025)	(0.0013)	(0.0060)	(0.0057)
	(0.0157)	(0.0154)	(0.0131)	(0.0130)
Husband is black	0.0039	0.0051	0.0033	0.0027
	(0.0207)	(0.0204)	(0.0190)	(0.0191)
Wife is white	0.0217	0.0220	0.0166	0.0169
	(0.0146)	(0.0144)	(0.0126)	(0.0125)
Wife is black	0.0258	0.0257	0.0249	0.0264
	(0.0239)	(0.0237)	(0.0208)	(0.0208)
Spouses are same race	(0.0168)	(0.0175)	(0.0116)	(0.0131)
	(0.0138)	(0.0136)	(0.0114)	(0.0114)
Head's Parents Intact	(0.0140)	(0.0140)	(0.0125)	(0.0126)
	(0.0064)**	(0.0063)**	(0.0057)**	(0.0056)**
Years Education Head	0.0071	0.0066	0.0059	0.0059
	(0.0054)	(0.0053)	(0.0042)	(0.0041)
Years Education Spouse	0.0091	0.0092	0.0079	0.0072
	(0.0056)	(0.0055)*	(0.0043)*	(0.0042)*
|Years Education Gap|	(0.0013)	(0.0015)	(0.0005)	(0.0008)
	(0.0018)	(0.0018)	(0.0016)	(0.0016)
Educ.*Educ.(Spouse)	(0.0006)	(0.0006)	(0.0006)	(0.0005)
	(0.0004)	(0.0004)	(0.0003)*	(0.0003)*
Male Spouse's earnings	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
Female Spouse's earnings	0.0001	0.0001	0.0002	0.0001
	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Female High Earner	0.0017	0.0018	0.0012	0.0010
	(0.0078)	(0.0078)	(0.0070)	(0.0070)
City Pop. >=500k	(0.0052)	(0.0056)	(0.0013)	(0.0011)
	(0.0105)	(0.0104)	(0.0091)	(0.0091)
100k<City Pop.<500k	0.0005	0.0000	(0.0010)	(0.0020)
	(0.0086)	(0.0085)	(0.0076)	(0.0077)
50k<City Pop.<100k	(0.0005)	(0.0021)	(0.0002)	0.0003
	(0.0110)	(0.0110)	(0.0100)	(0.0099)

25k<City Pop.<50k	(0.0141)	(0.0155)	(0.0147)	(0.0146)
	(0.0110)	(0.0110)	(0.0099)	(0.0099)
10k<City Pop.<25k	(0.0101)	(0.0114)	(0.0067)	(0.0059)
	(0.0095)	(0.0095)	(0.0083)	(0.0083)
1 Kid	0.0147	0.0149	0.0096	0.0112
	(0.0077)*	(0.0077)*	(0.0071)	(0.0071)
2 Kids	0.0087	0.0081	0.0045	0.0064
	(0.0078)	(0.0078)	(0.0070)	(0.0070)
3 Kids	0.0096	0.0108	0.0085	0.0079
	(0.0111)	(0.0108)	(0.0097)	(0.0097)
4 Kids	0.0332	0.0319	0.0169	0.0169
	(0.0168)**	(0.0168)*	(0.0141)	(0.0143)
5 Kids or More	0.0334	0.0339	0.0196	0.0167
	(0.0331)	(0.0318)	(0.0241)	(0.0244)
Owns House	(0.0096)	(0.0102)	(0.0150)	(0.0143)
	(0.0075)	(0.0075)	(0.0067)**	(0.0067)**
Have Debt	0.0025	0.0029	0.0042	0.0046
	(0.0062)	(0.0062)	(0.0055)	(0.0055)
IRA or Annuity	(0.0121)	(0.0127)	(0.0100)	(0.0092)
	(0.0091)	(0.0091)	(0.0081)	(0.0080)
Exactly one spouse smokes	0.0118	0.0125	0.0099	0.0099
	(0.0071)*	(0.0070)*	(0.0063)	(0.0063)
Moved last year	0.0008	0.0005	(0.0016)	(0.0012)
	(0.0065)	(0.0065)	(0.0061)	(0.0061)
Head Married > Once (=1)	0.0150	0.0164	0.0236	0.0234
	(0.0084)*	(0.0084)**	(0.0072)***	(0.0071)***
Same Industry (=1)	0.0086	0.0055	0.0036	0.0039
	(0.0096)	(0.0098)	(0.0096)	(0.0096)
Husband's Occupation's Earnings Variance	(0.0016)	(0.0016)	(0.0009)	(0.0010)
	(0.0017)	(0.0018)	(0.0015)	(0.0015)
Wife's Occupation's Earnings Variance	(0.0028)	(0.0038)	(0.0046)	(0.0038)
	(0.0025)	(0.0027)	(0.0023)**	(0.0025)
Spouses' Occupations' Earnings Covariance	(0.0384)	(0.0394)	(0.0424)	(0.0417)
	(0.0209)*	(0.0208)*	(0.0196)**	(0.0193)**
Unobserved Distance	Observed Only	Observed Only	First Full Time	Probabilistic
Household-Year Pairs	4141	4141	5211	5213
Includes Controls	Yes	Yes	Yes	Yes
Occupation Indicators	Yes	Yes	Yes	Yes
Log Likelihood	-629.43	-622.38	-771.48	-772.6
Pseudo R Squared	0.1046	0.1146	0.1203	0.1159

* p<0.1; ** p<0.05; *** p<0.01. This table contains the estimated marginal effects for the covariates in the divorce model. Columns 1-4 correspond, respectively, to Table 10 (Columns 1 and 2) and Table 12 (Columns 1 and 2). All standard errors are cluster robust.

Appendix B: Wage Correlation Decomposition

The de-meaned wages:

$$(B1)\ \widetilde{w}_{it} \equiv w_{it} - \overline{w}_t = \alpha_i - E^i(\alpha_i) + t\left(\beta_i - E^i(\beta_i)\right) + \varepsilon_{it} - 0,$$

are expressed as deviations from their time series expectations:

$$(B2)\ E^t(\widetilde{w}_{it}) = a_i - E^i(a_i) + E^t(t)\left(\beta_i - E^i(\beta_i)\right) + 0;$$

$$(B3)\ \dot{w}_{it} \equiv \widetilde{w}_{it} - E^t(\widetilde{w}_{it}) = \left(t - E^t(t)\right)\left(\beta_i - E^i(\beta_i)\right) + \varepsilon_{it}.$$

Covariance between two occupations (*i* and *j*) is defined:

$$(B4)\ \sigma_{ij} \equiv E^t\left[\dot{w}_{it}\dot{w}_{jt}\right] = E^t\left[\left(\widetilde{w}_{it} - E^t(\widetilde{w}_{it})\right)\ \left(\widetilde{w}_{jt} - E^t\left(\widetilde{w}_{jt}\right)\right)\right].$$

$$(B5)\ \sigma_{ij} = \dot{b}_i\dot{b}_j\sigma_t^2 + Cov\left(\varepsilon_{it}, \varepsilon_{jt}\right) + 0\ + 0, \text{where}$$

$$\dot{b}_i \equiv \beta_i - E^i(\beta_i).$$

The only terms that have a non-zero expectation (in B4) are the first two "diagonals", which have the interpretations, respectively, of "correlation in time trends" and "correlation in shocks".

The occupation-specific time trends are estimated from a random trends model. To estimate the occupation-specific trends, I use a method described in Wooldridge (2002). I take the first difference of (B1); this negates the fixed effect, "alpha i", but the trend ($\beta_i t - \beta_i(t-1) = \beta_i$) is now a fixed effect in the differenced model. I then estimate "beta i" using a fixed effects regression of change in average earnings on the transformed year-fixed effects. Using Wooldridge's notation, these are *xi* subscript t:

$$(\text{B6})\ \Delta w_{it} = \xi_t + \beta_i + \Delta\varepsilon_{it},$$

where the deltas represent first differences. The fixed component of the residuals can then be estimated by fitting the model, and these are the occupation-specific time trend estimates.[1] The time trend component of earnings correlation is the product of the two occupations' time trends (expressed as deviations from the mean) times a positive constant reflecting the length of the time series. This component is positive if both occupations' earnings trend faster than average or both trend slower than average and are negative otherwise.

1 This method is equivalent to (cross-sectionally) de-meaning the observations and regressing de-meaned earnings on time.

Appendix C: Earnings Correlation Model with Single Dimension Distances

Table C1: Estimates from regression of earnings correlation on single dimension distances.

Distance Measure	Log of Transformed Earnings Correlation, Occupations i and j	Correlation of Earnings Occupations i and j	Earnings Correlation Originating from Time Trends	Earnings Correlation Originating from Shocks
Arm-Hand Steadiness	0.1878	0.0799	0.0945	-0.0146
	(0.0515)***	(0.0228)***	(0.0253)***	(0.0195)
Auditory Attention	-0.0271	-0.017	0.0083	-0.0252
	(0.0346)	(0.0153)	(0.0170)	(0.0131)*
Category Flexibility	-0.0954	-0.0389	-0.0471	0.0082
	(0.0376)**	(0.0167)**	(0.0184)**	(0.0143)
Control Precision	-0.2214	-0.0942	-0.0557	-0.0385
	(0.0598)***	(0.0265)***	(0.0294)*	(0.0227)*
Deductive Reasoning	-0.2234	-0.1072	-0.1181	0.0109
	(0.0758)***	(0.0336)***	(0.0372)***	(0.0287)
Depth Perception	0.0253	0.0132	0.0138	-0.0006
	(0.0490)	(0.0217)	(0.0240)	(0.0186)
Dynamic Flexibility	0.1586	0.0681	0.0498	0.0182
	(0.0285)***	(0.0126)***	(0.0140)***	(0.0108)*
Dynamic Strength	0.2354	0.1103	0.1119	-0.0016
	(0.0673)***	(0.0299)***	(0.0331)***	(0.0255)
Explosive Strength	-0.1573	-0.0658	-0.0892	0.0234
	(0.0264)***	(0.0117)***	(0.0129)***	(0.0100)**
Extent Flexibility	0.0045	0.0022	-0.0192	0.0215
	(0.0582)	(0.0258)	(0.0286)	(0.0221)
Far Vision	-0.0449	-0.0211	-0.0154	-0.0057
	(0.0253)*	(0.0112)*	(0.0124)	(0.0096)
Finger Dexterity	-0.1841	-0.0836	-0.0825	-0.0011
	(0.0404)***	(0.0179)***	(0.0198)***	(0.0153)
Flexibility of Closure	0.0239	0.0115	-0.0021	0.0136
	(0.0311)	(0.0138)	(0.0152)	(0.0118)
Fluency of Ideas	-0.0313	-0.0196	0.0328	-0.0523
	(0.0696)	(0.0309)	(0.0342)	(0.0264)**
Glare Sensitivity	-0.3457	-0.1445	-0.1485	0.004
	(0.0476)***	(0.0211)***	(0.0234)***	(0.0181)
Gross Body Coordination	-0.0507	-0.0281	-0.0454	0.0173
	(0.0691)	(0.0306)	(0.0339)	(0.0262)
Gross Body Equilibrium	0.086	0.0331	0.0361	-0.0031

	(0.0472)*	(0.0209)	(0.0232)	(0.0179)
Hearing Sensitivity	0.1733	0.0772	0.0634	0.0139
	(0.0370)***	(0.0164)***	(0.0182)***	(0.0141)
Inductive Reasoning	0.4404	0.1946	0.1638	0.0308
	(0.0669)***	(0.0297)***	(0.0329)***	(0.0254)
Information Ordering	0.1181	0.052	0.048	0.0041
	(0.0432)***	(0.0192)***	(0.0212)**	(0.0164)
Manual Dexterity	0.212	0.0969	0.065	0.0319
	(0.0588)***	(0.0261)***	(0.0289)**	(0.0223)
Mathematical Reasoning	-0.1719	-0.0819	-0.0802	-0.0017
	(0.0588)***	(0.0261)***	(0.0289)***	(0.0223)
Memorization	0.0245	0.0096	0.009	0.0006
	(0.0334)	(0.0148)	(0.0164)	(0.0127)
Multilimb Coordination	0.1787	0.0734	0.0877	-0.0143
	(0.0647)***	(0.0287)**	(0.0318)***	(0.0245)
Near Vision	-0.0334	-0.0182	-0.0333	0.0151
	(0.0287)	(0.0127)	(0.0141)**	(0.0109)
Night Vision	0.221	0.0975	0.0853	0.0122
	(0.0536)***	(0.0238)***	(0.0263)***	(0.0203)
Number Facility	-0.0033	0.0019	-0.0228	0.0246
	(0.0527)	(0.0234)	(0.0259)	(0.0200)
Oral Comprehension	-0.2128	-0.0918	-0.0946	0.0028
	(0.0615)***	(0.0273)***	(0.0302)***	(0.0233)
Oral Expression	-0.0884	-0.0498	-0.0201	-0.0297
	(0.0590)	(0.0261)*	(0.0290)	(0.0224)
Originality	-0.0155	0.0044	-0.0231	0.0275
	(0.0658)	(0.0292)	(0.0323)	(0.0250)
Perceptual Speed	0.024	0.01	0.0099	0.0001
	(0.0291)	(0.0129)	(0.0143)	(0.0111)
Peripheral Vision	0.0736	0.0299	0.035	-0.0051
	(0.0591)	(0.0262)	(0.0290)	(0.0224)
Problem Sensitivity	-0.1473	-0.0676	-0.0402	-0.0274
	(0.0454)***	(0.0201)***	(0.0223)*	(0.0172)
Rate Control	-0.0796	-0.0371	-0.0456	0.0085
	(0.0637)	(0.0282)	(0.0313)	(0.0242)
Reaction Time	-0.3516	-0.1564	-0.153	-0.0034
	(0.0745)***	(0.0330)***	(0.0366)***	(0.0283)
Response Orientation	0.3226	0.1334	0.1057	0.0277
	(0.0662)***	(0.0293)***	(0.0325)***	(0.0251)
Selective Attention	0.0658	0.0315	0.029	0.0025
	(0.0273)**	(0.0121)***	(0.0134)**	(0.0104)

Sound Localization	-0.1723	-0.0731	-0.0797	0.0066
	(0.0479)***	(0.0212)***	(0.0235)***	(0.0182)
Spatial Orientation	0.0073	0.0007	0.023	-0.0223
	(0.0439)	(0.0195)	(0.0216)	(0.0167)
Speech Clarity	0.078	0.0375	0.0579	-0.0204
	(0.0472)*	(0.0209)*	(0.0232)**	(0.0179)
Speech Recognition	-0.0358	-0.0158	-0.0552	0.0394
	(0.0411)	(0.0182)	(0.0202)***	(0.0156)**
Speed of Closure	-0.0083	-0.0054	0.0025	-0.0079
	(0.0338)	(0.0150)	(0.0166)	(0.0128)
Speed of Limb Movement	0.0132	0.0079	0.0163	-0.0084
	(0.0470)	(0.0208)	(0.0231)	(0.0178)
Stamina	0.1796	0.0733	0.0956	-0.0222
	(0.0639)***	(0.0283)***	(0.0314)***	(0.0242)
Static Strength	-0.3381	-0.149	-0.1185	-0.0304
	(0.0642)***	(0.0285)***	(0.0315)***	(0.0243)
Time Sharing	0.0538	0.0265	0.024	0.0025
	(0.0247)**	(0.0110)**	(0.0121)**	(0.0094)
Trunk Strength	0.0119	0.0076	0.0112	-0.0036
	(0.0508)	(0.0225)	(0.0250)	(0.0193)
Visual Color Discrimination	0.0503	0.0191	0.0131	0.0059
	(0.0305)*	(0.0135)	(0.0150)	(0.0116)
Visualization	-0.0337	-0.0146	0.0064	-0.021
	(0.0331)	(0.0147)	(0.0163)	(0.0126)*
Wrist-Finger Speed	0.0729	0.0342	0.0342	0
	(0.0377)*	(0.0167)**	(0.0185)*	(0.0143)
Written Comprehension	0.0332	0.0193	0.0166	0.0027
	(0.0739)	(0.0328)	(0.0363)	(0.0281)
Written Expression	-0.1373	-0.0595	-0.0973	0.0378
	(0.0717)*	(0.0318)*	(0.0352)***	(0.0272)
Analyzing Data or Information	-0.2076	-0.0936	-0.0363	-0.0573
	(0.0501)***	(0.0222)***	(0.0246)	(0.0190)***
Assisting and Caring for Others	-0.0147	-0.0075	-0.02	0.0125
	(0.0232)	(0.0103)	(0.0114)*	(0.0088)
Coaching and Developing Others	0.047	0.0201	-0.0099	0.03
	(0.0410)	(0.0182)	(0.0202)	(0.0156)*
Communicating with Persons Outside	0.1163	0.0491	0.0484	0.0007

Organization				
	(0.0428)***	(0.0190)***	(0.0210)**	(0.0162)
Communicating with Supervisors, Peers, or Subordinates	-0.0009	-0.0032	-0.0171	0.0139
	(0.0391)	(0.0174)	(0.0192)	(0.0148)
Controlling Machines and Processes	-0.0864	-0.0379	-0.0454	0.0075
	(0.0468)*	(0.0207)*	(0.0230)**	(0.0177)
Coordinating the Work and Activities of Others	0.0577	0.0242	0.0218	0.0024
	(0.0387)	(0.0171)	(0.0190)	(0.0147)
Developing Objectives and Strategies	0.1544	0.0681	0.043	0.0251
	(0.0412)***	(0.0183)***	(0.0203)**	(0.0157)
Developing and Building Teams	0.0635	0.0284	0.0436	-0.0151
	(0.0383)*	(0.0170)*	(0.0188)**	(0.0145)
Documenting/Recording Information	-0.0819	-0.0386	-0.043	0.0044
	(0.0357)**	(0.0158)**	(0.0175)**	(0.0136)
Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment	0.0544	0.0241	0.0101	0.014
	(0.0343)	(0.0152)	(0.0169)	(0.0130)
Establishing and Maintaining Interpersonal Relationships	-0.0507	-0.022	-0.0108	-0.0112
	(0.0362)	(0.0161)	(0.0178)	(0.0137)
Estimating the Quantifiable Characteristics of Products, Events, or Information	0.0321	0.0149	0.0145	0.0004
	(0.0287)	(0.0127)	(0.0141)	(0.0109)
Evaluating Information to Determine Compliance with Standards	-0.0449	-0.0186	-0.023	0.0043
	(0.0318)	(0.0141)	(0.0156)	(0.0121)
Getting Information	-0.0703	-0.0306	-0.0432	0.0126
	(0.0383)*	(0.0170)*	(0.0188)**	(0.0145)
Guiding, Directing, and Motivating Subordinates	-0.2295	-0.1078	-0.079	-0.0288

	(0.0409)***	(0.0182)***	(0.0201)***	(0.0155)*
Handling and Moving Objects	-0.1124	-0.0385	-0.0673	0.0287
	(0.0455)**	(0.0202)*	(0.0223)***	(0.0173)*
Identifying Objects, Actions, and Events	-0.0093	-0.0047	-0.006	0.0013
	(0.0311)	(0.0138)	(0.0153)	(0.0118)
Inspecting Equipment, Structures, or Material	-0.09	-0.0458	-0.0609	0.0152
	(0.0369)**	(0.0164)***	(0.0181)***	(0.0140)
Interacting With Computers	0.1244	0.0578	0.0385	0.0193
	(0.0417)***	(0.0185)***	(0.0205)*	(0.0158)
Interpreting the Meaning of Information for Others	-0.1352	-0.0603	-0.0439	-0.0164
	(0.0417)***	(0.0185)***	(0.0205)**	(0.0158)
Judging the Qualities of Things, Services, or People	-0.0257	-0.0113	0.0075	-0.0188
	(0.0300)	(0.0133)	(0.0147)	(0.0114)*
Making Decisions and Solving Problems	0.0121	0.0017	-0.0153	0.017
	(0.0421)	(0.0187)	(0.0207)	(0.0160)
Monitor Processes, Materials, or Surroundings	-0.0603	-0.028	-0.0313	0.0033
	(0.0288)**	(0.0128)**	(0.0142)**	(0.0109)
Monitoring and Controlling Resources	0.0523	0.023	0.0101	0.0129
	(0.0331)	(0.0147)	(0.0163)	(0.0126)
Operating Vehicles, Mechanized Devices, or Equipment	-0.3087	-0.1395	-0.138	-0.0015
	(0.0446)***	(0.0198)***	(0.0219)***	(0.0169)
Organizing, Planning, and Prioritizing Work	0.008	0.0031	0.0052	-0.0021
	(0.0424)	(0.0188)	(0.0208)	(0.0161)
Performing Administrative Activities	-0.137	-0.0617	-0.0358	-0.0259
	(0.0372)***	(0.0165)***	(0.0182)**	(0.0141)*
Performing General Physical Activities	-0.1066	-0.0455	-0.0111	-0.0343
	(0.0512)**	(0.0227)**	(0.0251)	(0.0194)*
Performing for or Working Directly with the	0.0178	0.0035	-0.0115	0.015

Public				
	(0.0286)	(0.0127)	(0.0140)	(0.0108)
Processing Information	0.0005	-0.0001	-0.0124	0.0123
	(0.0451)	(0.0200)	(0.0221)	(0.0171)
Provide Consultation and Advice to Others	-0.0937	-0.039	-0.0348	-0.0043
	(0.0412)**	(0.0183)**	(0.0203)*	(0.0156)
Repairing and Maintaining Electronic Equipment	-0.1102	-0.0451	-0.0551	0.0101
	(0.0340)***	(0.0151)***	(0.0167)***	(0.0129)
Repairing and Maintaining Mechanical Equipment	0.1024	0.0358	0.0588	-0.023
	(0.0516)**	(0.0229)	(0.0254)**	(0.0196)
Resolving Conflicts and Negotiating with Others	0.0949	0.0445	0.042	0.0026
	(0.0386)**	(0.0171)***	(0.0189)**	(0.0146)
Scheduling Work and Activities	0.1154	0.0511	0.0304	0.0207
	(0.0403)***	(0.0179)***	(0.0198)	(0.0153)
Selling or Influencing Others	-0.0569	-0.0236	-0.0214	-0.0022
	(0.0331)*	(0.0147)	(0.0162)	(0.0126)
Staffing Organizational Units	0.1279	0.0601	0.0346	0.0255
	(0.0337)***	(0.0149)***	(0.0165)**	(0.0128)**
Thinking Creatively	-0.1287	-0.0574	-0.0498	-0.0076
	(0.0384)***	(0.0170)***	(0.0189)***	(0.0146)
Training and Teaching Others	0.0523	0.0279	0.0423	-0.0144
	(0.0370)	(0.0164)*	(0.0182)**	(0.0140)
Updating and Using Relevant Knowledge	0.0961	0.0484	0.0266	0.0218
	(0.0430)**	(0.0191)**	(0.0211)	(0.0163)
Active Learning	0.2815	0.1238	0.1576	-0.0339
	(0.0636)***	(0.0282)***	(0.0312)***	(0.0241)
Active Listening	-0.0496	-0.0225	-0.013	-0.0095
	(0.0652)	(0.0289)	(0.0320)	(0.0247)
Complex Problem Solving	-0.153	-0.0653	-0.0313	-0.034
	(0.0569)***	(0.0252)***	(0.0279)	(0.0216)
Coordination	-0.0163	-0.0071	-0.0278	0.0207
	(0.0362)	(0.0161)	(0.0178)	(0.0137)
Critical Thinking	0.174	0.0806	0.0376	0.043
	(0.0598)***	(0.0265)***	(0.0294)	(0.0227)*

Equipment Maintenance	0.1852	0.0814	0.0808	0.0006
	(0.0618)***	(0.0274)***	(0.0303)***	(0.0234)
Equipment Selection	0.0903	0.0489	0.0577	-0.0088
	(0.0440)**	(0.0195)**	(0.0216)***	(0.0167)
Installation	-0.4675	-0.2083	-0.2084	0.0001
	(0.0302)***	(0.0134)***	(0.0148)***	(0.0115)
Instructing	0.035	0.0147	0.031	-0.0163
	(0.0525)	(0.0233)	(0.0258)	(0.0199)
Judgment and Decision Making	0.0508	0.0249	-0.0103	0.0351
	(0.0591)	(0.0262)	(0.0290)	(0.0224)
Learning Strategies	0.0743	0.0307	0.0446	-0.0139
	(0.0549)	(0.0243)	(0.0270)*	(0.0208)
Management of Financial Resources	-0.0417	-0.0225	-0.0338	0.0113
	(0.0372)	(0.0165)	(0.0183)*	(0.0141)
Management of Material Resources	-0.0181	-0.0059	-0.0014	-0.0045
	(0.0365)	(0.0162)	(0.0179)	(0.0138)
Management of Personnel Resources	0.058	0.0272	0.0197	0.0074
	(0.0461)	(0.0204)	(0.0226)	(0.0175)
Mathematics	0.0624	0.0305	0.0506	-0.0201
	(0.0449)	(0.0199)	(0.0221)**	(0.0171)
Monitoring	-0.1051	-0.042	-0.0615	0.0194
	(0.0458)**	(0.0203)**	(0.0225)***	(0.0174)
Negotiation	0.0615	0.0203	0.0258	-0.0055
	(0.0502)	(0.0222)	(0.0246)	(0.0190)
Operation Monitoring	-0.025	-0.0067	-0.0112	0.0045
	(0.0539)	(0.0239)	(0.0265)	(0.0204)
Operation and Control	0.3522	0.1599	0.1479	0.0121
	(0.0574)***	(0.0255)***	(0.0282)***	(0.0218)
Operations Analysis	-0.0909	-0.0405	-0.0588	0.0183
	(0.0314)***	(0.0139)***	(0.0154)***	(0.0119)
Persuasion	-0.0548	-0.0228	-0.0196	-0.0032
	(0.0497)	(0.0220)	(0.0244)	(0.0189)
Programming	-0.0318	-0.0134	-0.0141	0.0007
	(0.0302)	(0.0134)	(0.0148)	(0.0115)
Quality Control Analysis	0.119	0.0526	0.0356	0.017
	(0.0415)***	(0.0184)***	(0.0204)*	(0.0157)
Reading Comprehension	0.0724	0.0316	0.0031	0.0285
	(0.0742)	(0.0329)	(0.0364)	(0.0281)

Repairing	-0.0736	-0.0383	0.0121	-0.0505
	(0.0621)	(0.0275)	(0.0305)	(0.0235)**
Science	-0.0585	-0.0286	0.0261	-0.0547
	(0.0313)*	(0.0139)**	(0.0154)*	(0.0119)***
Service Orientation	-0.4694	-0.213	-0.2369	0.0238
	(0.0388)***	(0.0172)***	(0.0190)***	(0.0147)
Social Perceptiveness	0.0004	-0.0015	0.0428	-0.0443
	(0.0444)	(0.0197)	(0.0218)**	(0.0168)***
Speaking	0.2083	0.0925	0.0893	0.0032
	(0.0697)***	(0.0309)***	(0.0342)***	(0.0264)
Systems Analysis	0.0141	0.0003	0.0417	-0.0414
	(0.0629)	(0.0279)	(0.0309)	(0.0239)*
Systems Evaluation	0.0892	0.0442	0.012	0.0322
	(0.0654)	(0.0290)	(0.0321)	(0.0248)
Technology Design	0.1519	0.0669	0.0636	0.0034
	(0.0279)***	(0.0124)***	(0.0137)***	(0.0106)
Time Management	-0.0434	-0.018	0.0414	-0.0594
	(0.0480)	(0.0213)	(0.0236)*	(0.0182)***
Troubleshooting	-0.1415	-0.0578	-0.0619	0.004
	(0.0548)***	(0.0243)**	(0.0269)**	(0.0208)
Writing	-0.15	-0.0649	-0.0499	-0.015
	(0.0733)**	(0.0325)**	(0.0360)	(0.0278)
Administration and Management	-0.0492	-0.0233	-0.0348	0.0114
	(0.0361)	(0.0160)	(0.0177)**	(0.0137)
Biology	-0.174	-0.0797	-0.0817	0.002
	(0.0264)***	(0.0117)***	(0.0129)***	(0.0100)
Building and Construction	0.2068	0.0958	0.0912	0.0046
	(0.0294)***	(0.0130)***	(0.0144)***	(0.0112)
Chemistry	0.0341	0.0179	0.0317	-0.0138
	(0.0286)	(0.0127)	(0.0140)**	(0.0109)
Clerical	0.0441	0.0187	0.0101	0.0086
	(0.0349)	(0.0155)	(0.0171)	(0.0132)
Communications and Media	-0.1414	-0.0626	-0.0625	-0.0002
	(0.0390)***	(0.0173)***	(0.0191)***	(0.0148)
Computers and Electronics	0.0425	0.0227	0.0559	-0.0332
	(0.0425)	(0.0188)	(0.0209)***	(0.0161)**
Customer and Personal Service	0.1028	0.0475	0.0952	-0.0476
	(0.0322)***	(0.0143)***	(0.0158)***	(0.0122)***

Design	-0.1697	-0.0787	-0.0738	-0.005
	(0.0376)***	(0.0167)***	(0.0185)***	(0.0143)
Economics and Accounting	-0.1275	-0.0557	-0.0462	-0.0095
	(0.0306)***	(0.0136)***	(0.0151)***	(0.0116)
Education and Training	0.0562	0.0254	-0.0025	0.0279
	(0.0340)*	(0.0151)*	(0.0167)	(0.0129)**
Engineering and Technology	-0.0047	-0.0054	0.0003	-0.0057
	(0.0417)	(0.0185)	(0.0205)	(0.0158)
English Language	-0.0619	-0.027	-0.0374	0.0105
	(0.0405)	(0.0180)	(0.0199)*	(0.0154)
Fine Arts	0.0306	0.0128	0.0111	0.0017
	(0.0222)	(0.0098)	(0.0109)	(0.0084)
Food Production	-0.0024	-0.0021	0.0109	-0.0129
	(0.0212)	(0.0094)	(0.0104)	(0.0080)
Foreign Language	0.0664	0.0318	0.0332	-0.0014
	(0.0244)***	(0.0108)***	(0.0120)***	(0.0093)
Geography	0.1351	0.0629	0.0503	0.0126
	(0.0301)***	(0.0133)***	(0.0148)***	(0.0114)
History and Archeology	-0.033	-0.0142	-0.0321	0.0179
	(0.0278)	(0.0123)	(0.0136)**	(0.0105)*
Law and Government	-0.0108	-0.0045	0.0198	-0.0243
	(0.0336)	(0.0149)	(0.0165)	(0.0128)*
Mathematics	0.0466	0.0197	0.0163	0.0034
	(0.0321)	(0.0142)	(0.0158)	(0.0122)
Mechanical	-0.2357	-0.1017	-0.127	0.0253
	(0.0453)***	(0.0201)***	(0.0223)***	(0.0172)
Medicine and Dentistry	-0.0145	-0.0066	-0.0021	-0.0045
	(0.0245)	(0.0109)	(0.0120)	(0.0093)
Personnel and Human Resources	0.0721	0.0338	0.0215	0.0123
	(0.0334)**	(0.0148)**	(0.0164)	(0.0127)
Philosophy and Theology	0.0379	0.0148	0.034	-0.0192
	(0.0326)	(0.0144)	(0.0160)**	(0.0124)
Physics	0.0764	0.0343	0.0197	0.0146
	(0.0342)**	(0.0152)**	(0.0168)	(0.0130)
Production and Processing	-0.0771	-0.0373	-0.0012	-0.0361
	(0.0255)***	(0.0113)***	(0.0125)	(0.0097)***
Psychology	-0.0934	-0.0386	-0.0495	0.0109
	(0.0363)**	(0.0161)**	(0.0178)***	(0.0138)
Public Safety and Security	-0.0802	-0.0358	-0.0169	-0.0189

	(0.0262)***	(0.0116)***	(0.0128)	(0.0099)*
Sales and Marketing	0.0003	-0.0006	-0.0008	0.0002
	(0.0296)	(0.0131)	(0.0145)	(0.0112)
Sociology and Anthropology	0.0728	0.0308	0.0272	0.0036
	(0.0359)**	(0.0159)*	(0.0176)	(0.0136)
Telecommunications	0.0567	0.0242	0.0219	0.0023
	(0.0255)**	(0.0113)**	(0.0125)*	(0.0097)
Therapy and Counseling	0.003	0.001	0.0023	-0.0013
	(0.0313)	(0.0139)	(0.0154)	(0.0119)
Transportation	-0.0868	-0.0404	-0.052	0.0116
	(0.0255)***	(0.0113)***	(0.0125)***	(0.0097)
Industry Employment Shares	0.1185	0.0493	0.0017	0.0477
	(0.0696)*	(0.0309)	(0.0342)	(0.0264)*
Industry Shares Squared	-0.4814	-0.2099	-0.1348	-0.0751
	(0.1700)***	(0.0754)***	(0.0835)	(0.0645)
Occupation Shares Across Industries	-0.1835	-0.0818	-0.0798	-0.002
	(0.0398)***	(0.0176)***	(0.0195)***	(0.0151)
Occupation Shares Squared	0.0888	0.04	0.0396	0.0003
	(0.0268)***	(0.0119)***	(0.0132)***	(0.0102)
Constant	0.2789	0.1261	0.117	0.0092
	(0.0159)***	(0.0071)***	(0.0078)***	(0.0060)
R Squared	0.05	0.05	0.04	<0.01
Sample Size	70,876	70,876	70,876	70,876

* p<0.1; ** p<0.05; *** p<0.01.

CURRICULUM VITAE

Benjamin Van Kammen

Place of birth: Milwaukee, WI

Education

B.A., University of Wisconsin-Milwaukee, December 2007
Majors: Economics and History

Ph.D., University of Wisconsin-Milwaukee, August 2013

Dissertation Title: "Three Essays in Labor Economics"

Presentations

"Employment Effects of Paid Sick Leave Mandates."
Midwest Economics Association (MEA) meetings, 2012.
UW-Milwaukee Economics Seminar Series, 2011.
"Occupation Proximity and Marital Stability."
Midwest Economics Association (MEA) meetings, 2013.

Teaching

Econ 103: Principles of Microeconomics
Econ 210: Economic Statistics
Econ 301: Intermediate Microeconomics
Econ 415: Economics of Employment and Labor Relations
Econ 447: Labor Economics

Honors and Awards

Richard Perlman Paper Prize in Labor Economics, 2012.

Organizational Memberships and Service

American Economic Association 2010-Present
Midwest Economic Association 2011-Present
Chair and Discussant, MEA Meetings, 2012-2013.
Society of Labor Economists 2012-Present