Trade, Occupation Sorting, and Inequality *

WORK IN PROGRESS

Mons Chan[†] and Ming Xu[‡]

July 2017

Abstract

Firms react to changes in factor prices with intensive and extensive-margin employment adjustments at the occupational-level. We study the distributional and aggregate consequences of this make-or-buy dynamic by developing a novel network model of heterogeneous firm-to-firm trade where the boundary of each firm depends on factor prices and firm-occupation comparative advantage in input-production. We show that the model can be easily aggregated and taken to industry-level data, and use the calibrated model to examine recent trends in employment, wages and trade in the USA. We use public OES and CPS data to show empirical evidence that a significant fraction of the growth in wage inequality in the USA is due to changes in firm/industry specialization and occupation sorting. To understand and measure the underlying causes of these trends, we calibrate the model to occupation and industry data from the OES and input-output tables. The results suggest that 1/3rd of the increases in wage inequality stem from decreases in inter-industry trade frictions with the remaining 2/3rds stemming from changes in technology and labor supply. Falling trade frictions are also responsible for all of the increases in occupational sorting and concentration. Had trade frictions been held at their 2002 level, productivity growth would have led to an increase in vertical integration, rather than the decrease observed in the data.

*We thank Naoki Aizawa, Pol Antràs, Anmol Bhandari, Fatih Guvenen, Kyle Herkenhoff, Tom Holmes, Loukas Karabarbounis, Benjamin Malin, Ellen McGrattan, Amil Petrin, Victor Rios-Rull, Frederic Warzynski, Kei-Mu Yi and many seminar participants for their useful guidance, discussion and comments.

[†]University of Minnesota, chanx186@umn.edu

[‡]University of Minnesota,: xuxx0262@umn.edu

1 Introduction

A broad literature exists documenting and attempting to explain rising wage inequality in the USA and other nations. One significant feature of the change in wage inequality is that much of the increase appears to have manifested as between-firm rather than within-firm inequality. This has prompted speculation[1] that changes in the wage distribution are being driven by factors such as changes in the dispersion of firm-level productivity, or increased sorting of workers across firms and industries by skill. A third possibility, which we propose in this paper, is that changes in wage inequality are significantly affected by the increased sorting of occupations across firms and industries. This first has mechanical implications for between vs. within inequality, and secondly is also importantly linked to changing wages within and between occupation groups, changes in labor demand and overall efficiency.

Our main argument is as follows. Firms can either make or buy the inputs into their production process. This decision is influenced by the relative costs of inter-firm trade and in-house production (coordination costs vs. management or agency costs). As these costs decline[2], firms will increasingly choose to purchase inputs rather than make them in-house (ie: decreasing vertical integration or increasing specialization). This leads firms to hire fewer workers of that occupation type, and the firm from whom they purchase the input will likely hire more. On the extensive margin, this decreases the cardinality of the set of occupations hired by each firm[3], and on the intensive margin increases measures of occupation employment concentration. We call this increasing concentration of occupation employment share within firms/industries "Occupation Sorting".

Consider the following very simple example with two firms, A and B. Firm A employs 8 Lawyers and 2 Janitors. Firm B employs 2 Lawyers and 8 Janitors. Then suppose that one Janitor moves A $\rightarrow$ B, and one lawyer moves B $\rightarrow$ A. If everyone has the same wage, then there is no change in wage dispersion. However, if we suppose Lawyers at both firms make \$20 while Janitors at A make \$15 and Janitors at B make \$10, then this simple reshuffling of workers dramatically alters observed overall, within and between wage variance. In our example, overall wage variance increases from 22 to 24, between firm variance increases from 12 to 18, while within variance decreases from 10 to 6.

This effect can be clearly seen in the data. Figure 1 shows the evolution of industry-

[1]See Barth et al. (2014), and Faggio, Salvanes and Van Reenen (2010)

[2]See Holmes and Snider (2011) for a discussion of the evidence for declining outsourcing frictions

[3]Chan (2016) shows evidence of significant extensive-margin adjustment in response to trade and technology shocks for Denmark.

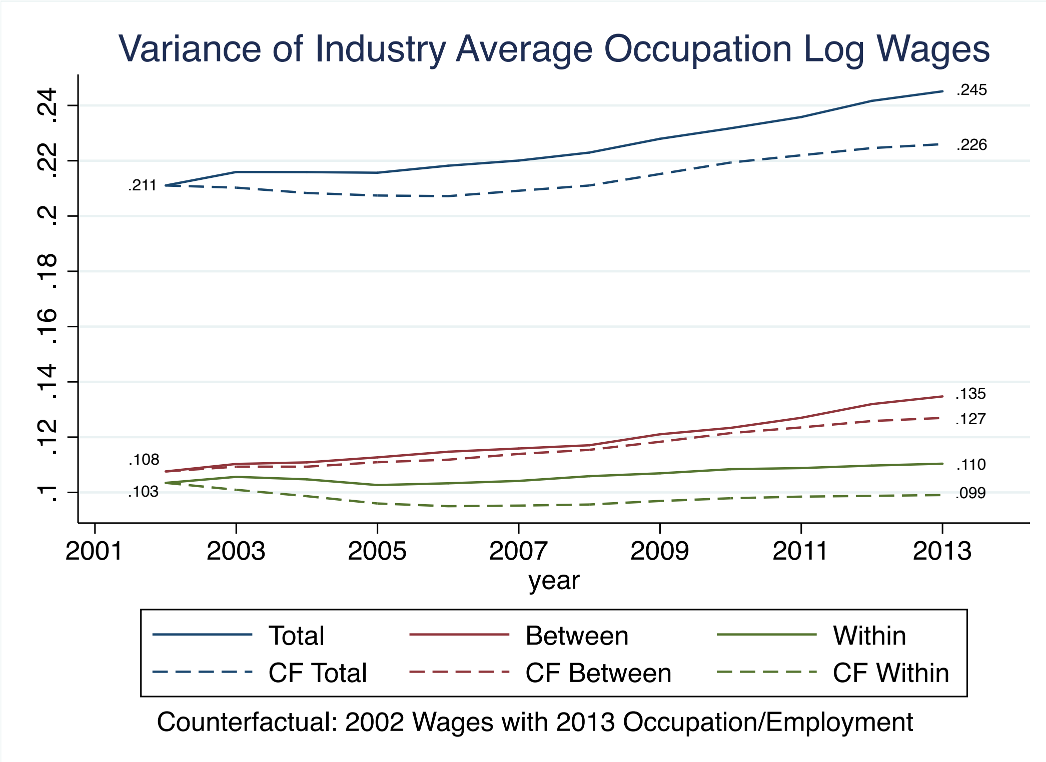

Figure 1

occupation average log wage variance from 2002 through 2013 calculated using OES public data [4] (the solid lines). It also shows the evolution of the counterfactual wage variance when we hold industry-occupation wages constant, and only allow the distribution of occupations across industries to change (dashed lines). The decomposition of both variances into within and between clearly shows that most of the increase in between variance is due to occupation sorting and shifts in labor demand, while this same effect decreased within variance, as in our simple example. This implies that changing occupational wages at the industry level contributed to rising wage inequality primarily via increased within-industry variance. We argue that this is also likely true at the firm and establishment level. Also note that we believe this effect has been relatively small since 2000. Our preliminary work with the Current Population Survey (Flood et al. (2016)) and PSID data suggest that the effects we propose and measure here would likely be much stronger through the 80s and 90s.

One question we want to address is how much of the change in observed between and within wage inequality is being driven simply by moving workers with different wages around

[4]See section 3 for a more thorough discussion of the data and methodology. This variance represents about 40% of total individual wage variance.

between firms. This is similar to the idea behind skill sorting, in that if workers of different skills have different average wages, and workers are somehow being increasingly grouped by skill at the firm level, then between-firm wage variance will increase. Indeed the two ideas overlap, since workers in different occupations will likely have different levels of skill. However, the mechanism is very different. Occupation sorting arises from firm-level production decisions which are relatively independent of worker skill. The fact that it may happen to coincide with some level of skill sorting is more of a side-product of the primary mechanism - firms are not choosing to sort on skill but rather on tasks. So, we would like to be able to separate how much of the changes in wages and sorting are due to occupational sorting vs. skill sorting. We also wish to investigate how much this mechanism - firm specialization and occupation sorting - has contributed to changing wages between and within occupations through changes in labor demand.

In order to investigate our proposed mechanism, we employ a model with heterogeneous firms and industries who choose to buy intermediate inputs or make them in-house by hiring from a set of heterogeneous occupations. Our model is different in that most of the rest of the literature on occupations, sorting and wages either assume a single type of output good, or index workers purely on skill[5]. We also differ from the trade literature in that firms select into multi-state outsourcing rather than exporting, resulting in complex vertically integrated linkages through the economy. Our model treats occupations as monopolistic providers of unique services or tasks in an environment with iceberg contracting/agency costs and direct linkages between industries, which we believe is unique in the literature[6].

One of our primary contributions is to document several facts in the data, and to build a tractable model of outsourcing and occupation sorting which is able to replicate those facts in a transparent manner. The first observation is that the growth in wage variance since about 2000 has been entirely driven by growth in the between industry and within industry, between occupation wage variance (section 2.2). The second observation is that this growth in the between industry, between occupation variance is driven by two effects: An increase in between industry wage variance due to compositional change and sorting, and an increase in within industry variance due primarily to changes in wages (2.2). The third observation is that occupational concentration and industry specialization has been increasing over the

[5]See Hagedorn, Law and Manovskii (2012) for a recent discussion of the literature on labour search/matching and occupational sorting

[6]Several recent papers examine models with networks of linkages using the Input-Output tables for the USA and other countries, including Kehoe, Ruhl and Steinberg (2013), De Soyres (2015) and Chaney (2014). However, none of these have an environment with endogenously heterogeneous wages across occupations and industries or with multiple outsourcing decisions

last decade (2.3). The fourth observation is that low skill occupations tend to experience significant wage declines as industries become more specialized, while high skill occupations experience wage growth (2.4).

Our second main contribution is to take our theoretical model of outsourcing and trade to US data on industry and occupation-level employment, wage and output using the OES and Input-Output Tables. We use the calibrated model to examine several trends in the data and run counter-factual exercises in order to decompose changes in wages, specialization and inequality. We find that about 1/3rd of the increase in wage inequality is due to decreases in inter-industry trade frictions, while the other 2/3rds are due to changes in technology and labor supply. These decreasing trade frictions have also driven all of the increases in concentration – if trade frictions had remained at their 2002 level, industry/occupation concentration would have decreased by 4.3% instead of increasing.

Our theoretical work is closely related to the recent literature on trade and inequality (Helpman et al., 2015; Grossman, Helpman and Kircher, 2015; Grossman and Helpman, 2015; Itskhoki and Helpman, 2015), and the literature on trade, tasks and firm organization (Caliendo and Rossi-Hansberg, 2012; Caliendo, Monte and Rossi-Hansberg, 2015; Grossman and Rossi-Hansberg, 2008).

We proceed by presenting our data and empirical results in section 2. We write down our theory of trade, sorting and outsourcing in section 3, followed in section 4 by our model calibration and counter-factual exercises. We also include in the appendices a simplified 2x2x2 version of our model which uses union wage bargaining and management costs to model the empirical trends we see in the data.

2 Data and Empirical Evidence

Our empirical methodology is employed with several goals in mind. First, we want to explore the implications of our simple model (see section 3) and verify whether the model predictions are consistent with the data. Second, we want to document to what extent these changes in sorting and specialization are occurring (if at all). Lastly, we want to tease out the empirical relationship between occupation sorting and the wage distribution.

We first outline the data we use in our empirical analysis and present evidence of increasing wage inequality in the USA, identifying which components are due to compositional vs.

wage changes. We then show evidence of increasing occupation sorting and specialization over the last decade. Finally we examine the link between wage changes and occupation concentration at the industry level.

2.1 Data and Methodology

For our initial analysis, we use public data from the Occupation Employment Statistics (OES) database, which includes data on occupation employment and wages at the disaggregated industry level. Specifically, each observation is a year-industry-occupation with information on occupation average yearly wages (as well as within-occupation dispersion) and employment within the industry. Due to measurement issues in the data, we restrict our analysis to 4-digit industries (300) and 2-digit occupations (23).

We restrict our analysis of the OES to the period between 2002 and 2013. This is primarily because the NAICS classification system changed in 2002 from classifying establishments by the industry of their parent firm to classifying them by their primary activity. For example, prior to 2002, Walmart's head office would have been classified as a retail store, while after 2002 it would have been included in the management services industry. Restricting our analysis to 2002-2013 allows a consistent measurement of industry employment concentration over time.

In order to frame the results we get from the OES, we also calculate wage variance using the March Current Population Surveys (IPUMS, Flood et al. (2016)) from 1982 through 2013. We clean the data following Lemieux (2006) and Autor, Katz and Kearney (2008)[7], providing us with wage variance measures consistent with the rest of the literature. Both the OES and CPS use occupation and industry classification systems which change over time[8]. Since we are concerned with measuring between and within industry/occupation variance, we employ crosswalks from the BEA and Census in order to carefully construct consistent measures of occupation and industry over our period of interest.

[7]Specifically, we restrict the sample to full time, full-year, non-government, non-farm wage workers ages 16 to 65. We exclude individuals with imputed wage observations and use appropriate sampling weights (multiplied by units of time) in all calculations. We follow Autor, Katz and Kearney (2008) in correcting for top and bottom-coded observations, and deflate wages using the chain-weighted GDP personal consumption expenditure deflator.

[8]See Autor and Dorn (2013) for a discussion of changes in occupation classifications over time

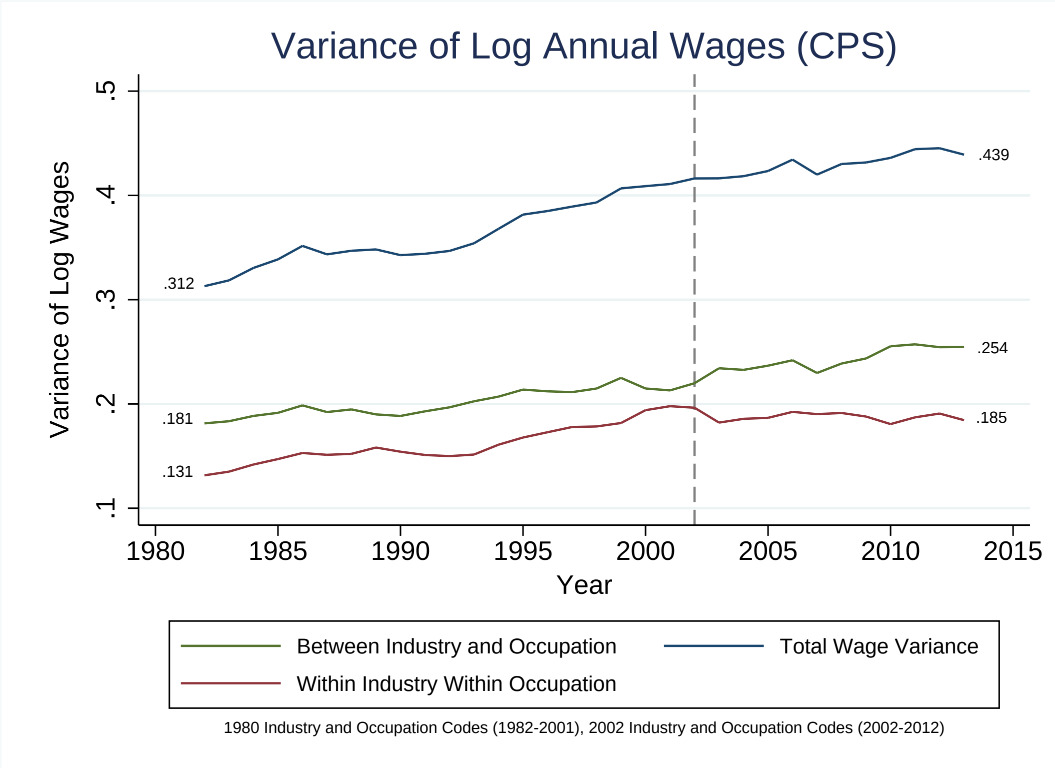

Figure 2

2.2 Trends in Wage Variance

The key motivation for our research is that wage inequality, as measured by the variance of log wages, has been increasing consistently over the last few decades. We first document this change using the CPS, compare it to what we see in the OES, then proceed with a decomposition exercise to separate the effects of composition change vs. changes in prices.

The blue line in figure (2) shows total variance of individual log annual wages between 1982 and 2013 for all full-time full-year workers (male and female)[9]. By this measure, wage inequality has grown by 40% since the early 1980s. We can decompose the total wage

[9]These results generally mirror those in Autor, Katz and Kearney (2008) and Lemieux (2006), though we use annual wages rather than weekly or hourly wages in order to coincide with the OES annual wage data.

variance into its component parts as follows:

$$\begin{aligned}
\text{Var of log wages} &= \mathbb{E}_{ion}[(w_{ion} - \bar{w})^2] && (1)\\
&= \mathbb{E}_{i}[(\bar{w}_i - \bar{w})^2] + \mathbb{E}_{ion}[(w_{ion} - \bar{w}_i)^2] && (2)\\
&= \mathbb{E}_{i}[(\bar{w}_i - \bar{w})^2] + \mathbb{E}_{io}[(\bar{w}_{io} - \bar{w}_i)^2] + \mathbb{E}_{ion}[(w_{ion} - \bar{w}_{io})^2] && (3)\\
&= \mathbb{E}_{io}[(\bar{w}_{io} - \bar{w})^2] + \mathbb{E}_{ion}[(w_{ion} - \bar{w}_{io})^2] && (4)
\end{aligned}$$

where w_{ion} is the log wage for individual n in occupation o and industry i. $\bar{w}$ is the overall mean wage, while $\bar{w}_i$ is the mean wage in industry i. Thus, the first term on the right in equation (3) is the "between industry" variance, ie: the variance of the mean industry wages around the overall mean wage. The second term is "within industry, between occupation" variance, and the last term is "within occupation" variance, which measures the variance of individual wages around the mean wage for their occupation in their industry of employment. The first term in equation (4) combines the within industry, and within industry between occupation component into the between industry, between occupation variance. Figure (2) includes this last decomposition, showing that well over half of all wage variance is between industries and between occupations (about 58% of total variance throughout this time period). Interestingly, the increase in overall variance between 1982 and 2000 is due equally to increases in between and within variance components, while the post-2000 increase is due entirely to increases in between variance[10].

Since we will use the OES for most of our subsequent empirical analysis, we want to verify that wage trends observed in the OES are comparable to wage trends in the CPS. The OES and CPS are fundamentally different data sets. The CPS is non-panel micro survey data created by randomly sampling individuals from the population[11], while the OES is a rolling survey of all non-farm US establishments[12]. In this sense, the OES is aggregate macro-level data and is better suited for detailed analysis of industry and occupation wage trends than the CPS, which in a given year may not have any observations for a given industry-occupation cell.

Figure (3) shows total wage variance in the CPS from 2002 through 2013, as well as

[10]The decompositions before 2002 and after 2002 are not necessarily directly comparable, as they are constructed using different industry and occupation classification code systems. However, the change in codes does not appear to dramatically alter measured wage variance.

[11]The CPS does actually have some panel structure, in that individual households are surveyed for several months in a row in one year, and then for another few months again the next year. However, the annual March CPS data we use does not include any panel structure.

[12]Every establishment is surveyed once every three years

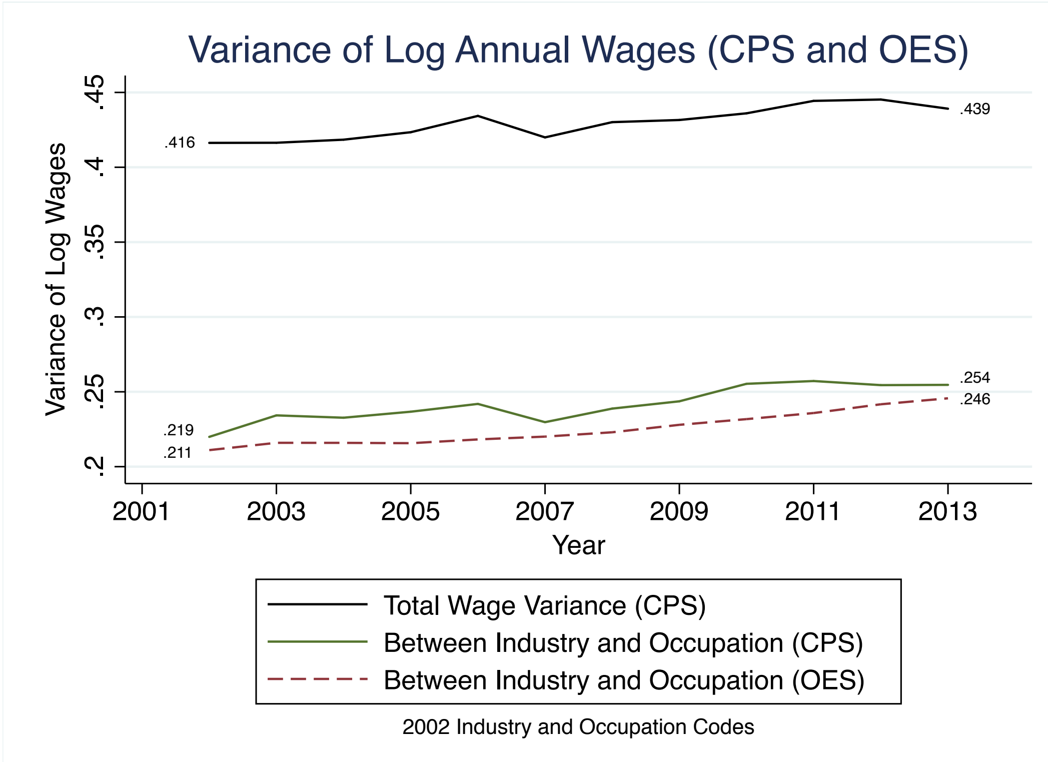

Figure 3

between industry, between occupation variance in both the CPS and the OES[13]. Remarkably, these two series are very close, giving us confidence that our analysis with the OES mirrors what we would see in other data sets, and that the between industry between occupation variance is an important component of overall wage inequality.

2.2.1 Counterfactual Wage Variance Decompositions

A direct method of measuring the contribution of occupation sorting to changes in wage inequality is to construct counter-factual measures of the variance by alternately holding fixed and allowing to change the share weights used in constructing the variance. Note that the variance of log industry-occupation wages is constructed using three separate objects - the wage observations, the within-industry weights associated with each wage observation, and the weight associated with each industry. We construct these weights using occupation employment shares. This allows us to construct counter-factual measures of changes in variance by, for example, holding the occupation employment shares fixed at their 2002

[13]Since the OES only has observations at the industry-occupation level, we can only see between industry and within industry, between occupation variance.

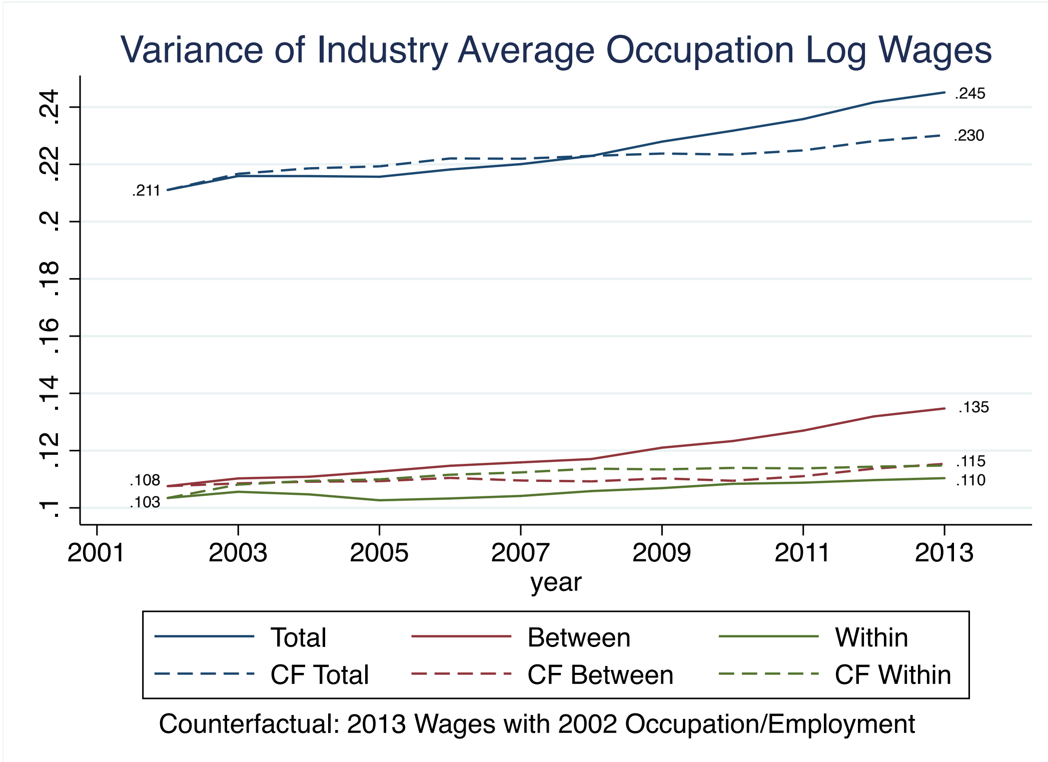

Figure 4

values, while allowing wages to evolve as they do in the subsequent years. The resulting trend in (counter-factual) variance illustrates the contribution of changes in the wage schedule to the changes in wage variance as compared to the contribution of sorting, which is measured by changes in occupation-industry employment shares. We construct both counter-factual cases (holding sorting fixed and allowing wages to evolve, and holding wages fixed while occupation shares evolve).

Figure 1 shows the factual changes in industry average occupation log wage variance (solid lines) as well as the first counter-factual, where we hold wages fixed and only allow occupation employment shares to change (the dashed lines). First, it is worth noting that even in our limited industry-occupation aggregate data, we see the trends in wage dispersion discussed in other papers. Overall wage variance is increasing, with increases in observed between variance accounting for almost all of the increase. However, our variance decomposition exercise shows that most of that increase in between industry variance is coming directly from the occupation sorting effect. The occupation sorting effect also acts to decrease within-industry wage variance during this period. This suggests that most of the contribution to growing wage variance from changes in actual occupation-level wages is coming through an

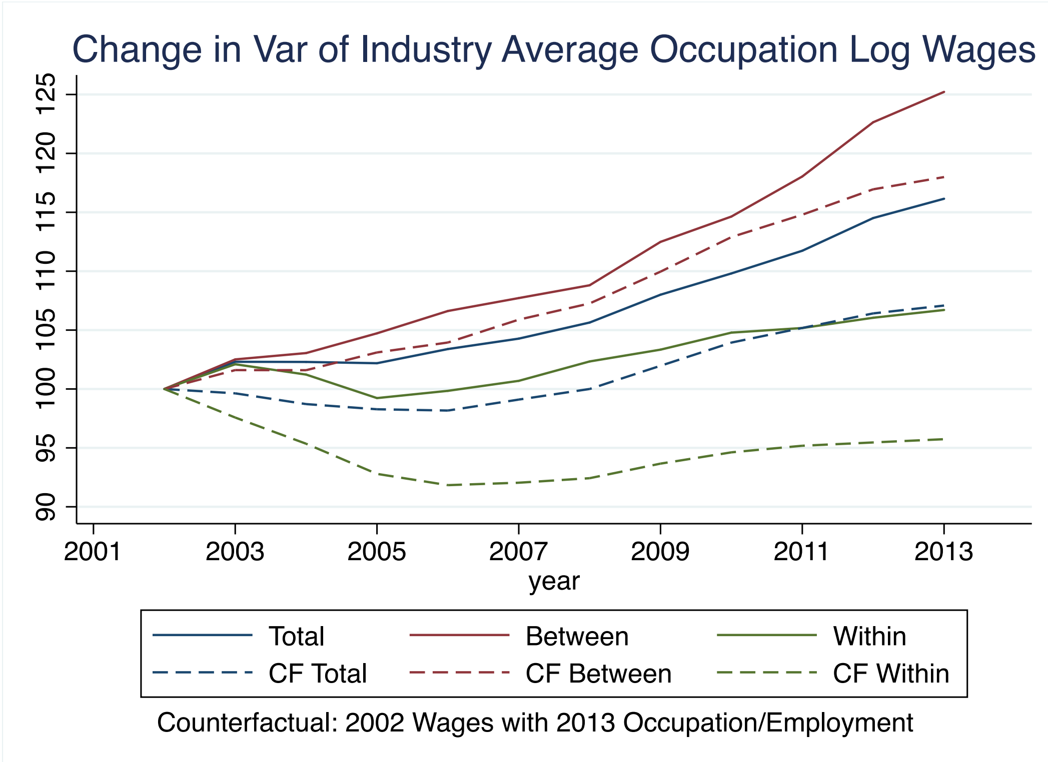

Figure 5

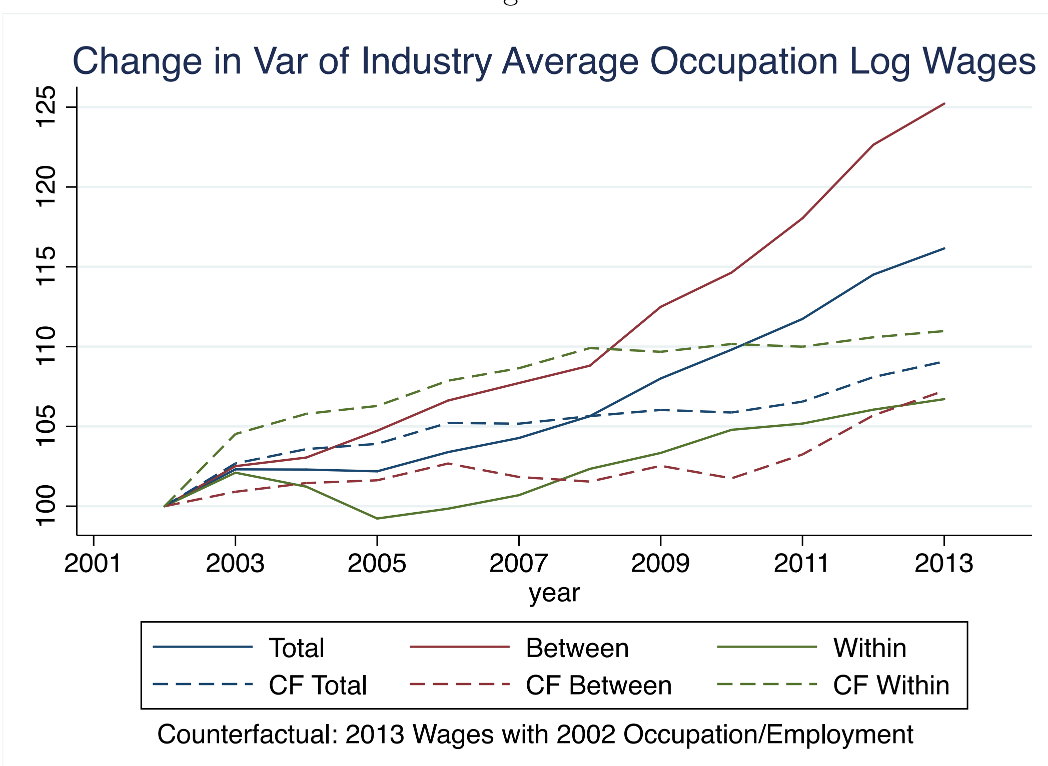

Figure 6

increase in within-industry inequality.

To see this more clearly, consider figure 4, which graphs the second counter-factual, where only wages are allowed to shift, and the distribution of occupations across and within industries is held fixed. It's clear that changing wages manifested primarily as a within-industry phenomenon, barely affecting between-industry variance at all. What's also interesting, is that changes in wage-levels accounted for much of the changes in overall wage variance between 2002 and 2008, but diverged significantly afterwards, suggesting an increasing influence from changes in occupation sorting. Figures **??** and 6 show the same numbers, with the first year normalized to 100.

This exercise makes it clear that the occupation sorting effect represents a significant driver of wage inequality, which is consistent with our model. In particular, the empirical results we have presented clearly show that occupation sorting and specialization is increasing in the United States, and that the interpretation of the relationship between changes in between/within inequality and overall inequality depends significantly on this sorting effect. As mentioned above, our results are likely a lower bound, as using industry-occupation level data such as the OES prevents us from observing much of the heterogeneity in occupation employment and movement between firms and establishments. Since our model is really a model of firm behaviour rather than industry behaviour, our results should be at this stage interpreted with caution. However, our ongoing research suggests that taking our analysis to firm-level data will give us similar results.

2.3 Trends in Specialization and Occupation Sorting

2.3.1 Previous Evidence

The observation that firms are becoming increasingly specialized with respect to intermediate production and employment is not new, though we believe our argument on how it is linked to wage inequality is novel. Yuskavage, Strassner and Medeiros (2008) show that purchased services share of gross output for all US industries grew from 22.5% in 1997 to 26% in 2006, while Baldwin, Beckstead and Caves (2001) document monotonically increasing firm specialization in Canada between 1975 and 1997. Other recent work has documented increasing occupational segregation at the firm and establishment level in both the USA (Handwerker and Spletzer, 2013) and West Germany (Card, Heining and Kline, 2013).

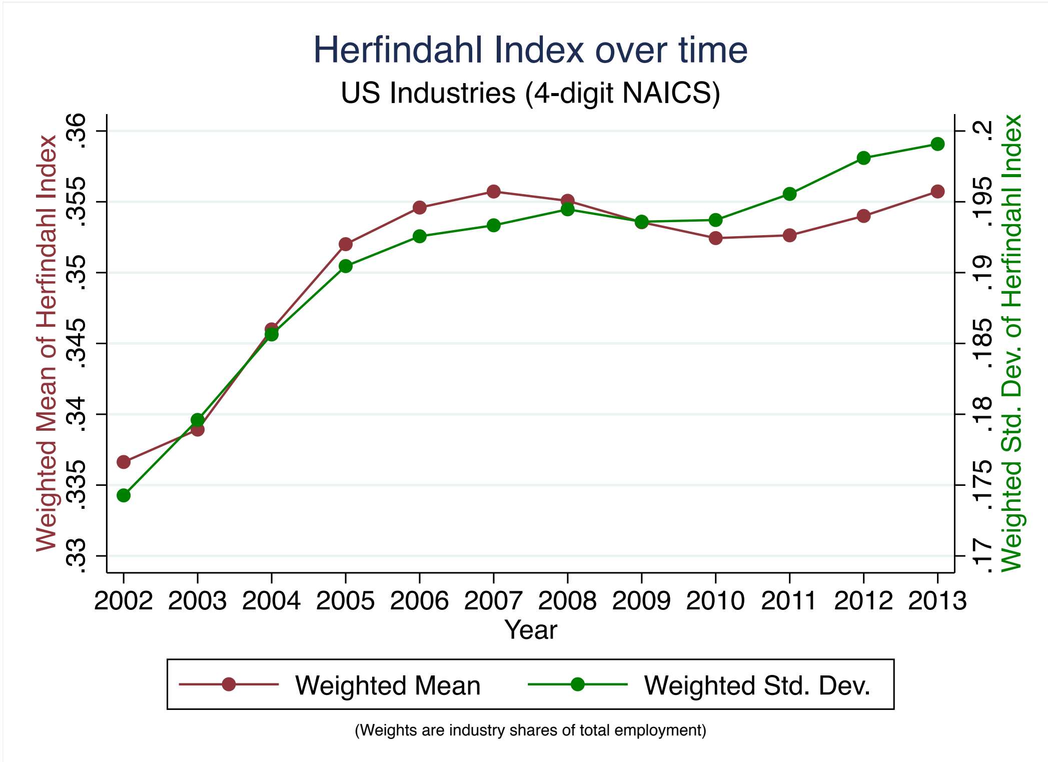

Figure 7

2.3.2 Changes in Occupation Concentration

Our results using the OES data is consistent with both this literature and the predictions of our model. Figure 7 shows the change in the employment-weighted mean Herfindahl index across all US industries between 2002 and 2013, as well as the change in the weighted standard deviation. There is a clear positive (statistically significant) trend for both dispersion in concentration, and average concentration, with a slight dip during the great recession. While the absolute change in Herfindahl seems small, it actually represents significant change, especially since it is a weighted average over all industries. The approximate change in average Herfindahl is from 0.335 to 0.355 over this period. To understand what this change means, consider the following two scenarios. First, suppose an industry employs three different occupations, each with 33% of employment. This generates a Herfindal of about 0.33. Now suppose one half of all workers in one of these occupations are outsourced (16% of total employment). This takes the Herfindal from 0.33 to 0.36. On the other extreme, suppose an industry employs all 22 2-digit occupations, with one occupation taking 57% of employment and the others each with ∼2%. If the industry outsources two of those occupations entirely (∼4% of employment), the Herfindal increases from 0.33 to 0.36 as well. So this

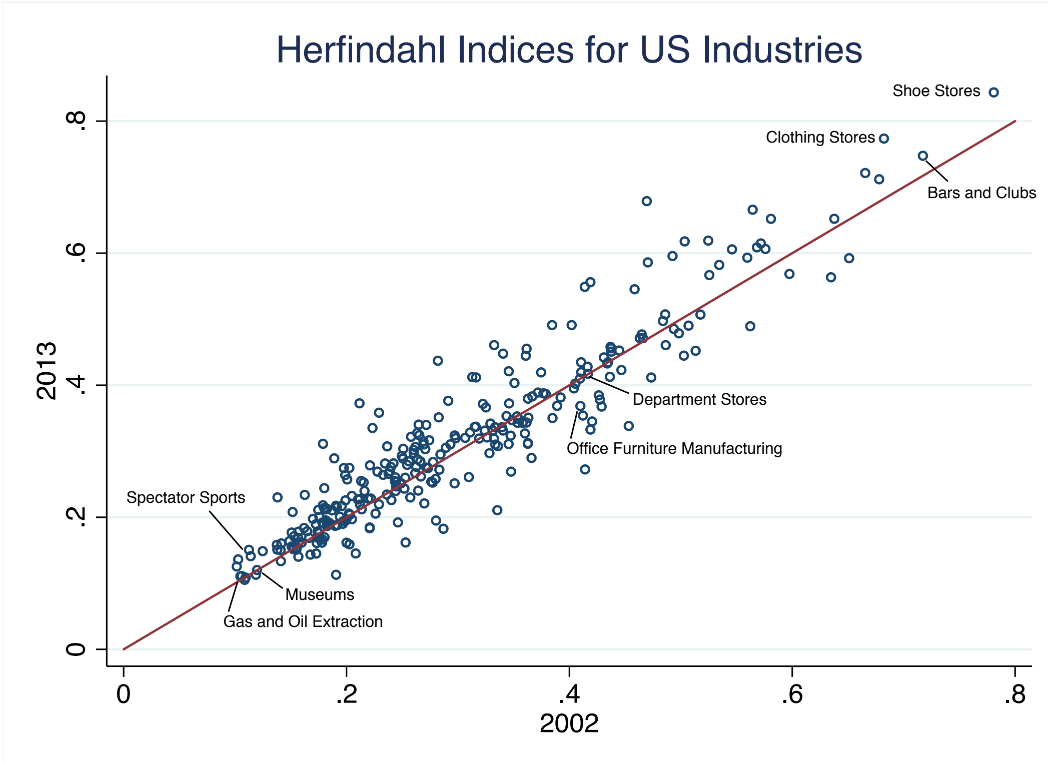

Figure 8

seemingly small change in the Herfindahl Index can be the result of a 4% to 16% change in total employment at the industry level, and dramatic change at the occupation level. This is consistent with the predictions of the model, assuming that outsourcing costs have been decreasing during this period.

The change in occupational employment concentrations across industries can be seen more clearly by focusing on the distribution of industries itself. Figure 8 plots the herfindahl for each industry in 2002 and 2013. Industries above the red 45-degree line increased employment concentration during this period. For example, the shoe store industry became more specialized, while the office furniture manufacturing industry became less specialized. The shoe store industry also happens to be the most concentrated (primarily consisting of workers in sales occupations) while the gas and oil extraction industry is highly diverse, employing a broad mix of many occupations.

Figure 9 shows each industry ranked by its change in employment herfindahl. Roughly 2/3rds of industries in the USA have become more occupationally concentrated during this period, with only 1/3rd becoming less specialized. In particular, note the liquor store industry, which has increased in employment herfindahl from 0.47 to 0.68. To get an idea

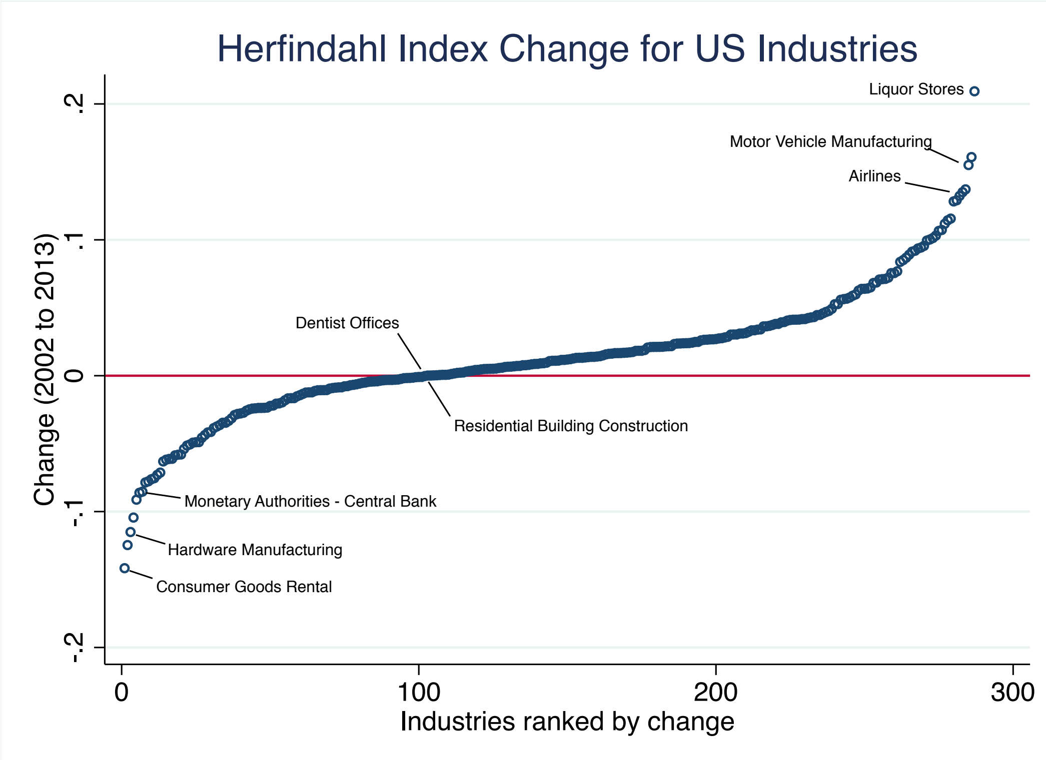

Figure 9

of what this change in occupation concentration means, between these years the share of transportation occupations in the liquor store industry went from 11% down to 2%, implying increased outsourcing of transportation services. Employment shares for management, food prep and office administration occupations also declined, while the employment share of sales and marketing occupations increased from 66% to 82%. Similarly, note the increase in the herfindahl for the airline industry (0.23 to 0.36). This was driven partially by a big decrease in employment in maintenance occupations (14% to 8%, a decrease of 38,000 employees), accompanied by an increase of employment of the same occupation in the support activities for air transportation industry, which is the primary industry for the airline maintenance occupation[14]. This is exactly the sort of occupational sorting predicted by our model, with intermediate occupations increasingly being sorted into their associated primary-output industries.

These shifts in occupational concentration are indicative of a general trend across US industries. Figures 10 and 11 show a general shift to the right in the distribution of the

[14]Part of this increase in airline industry employment herfindahl is also due to a change in occupation classifications within the airline industry. We are currently working on controlling for these irregularities, as well as obtaining more detailed data at the firm/establishment level.

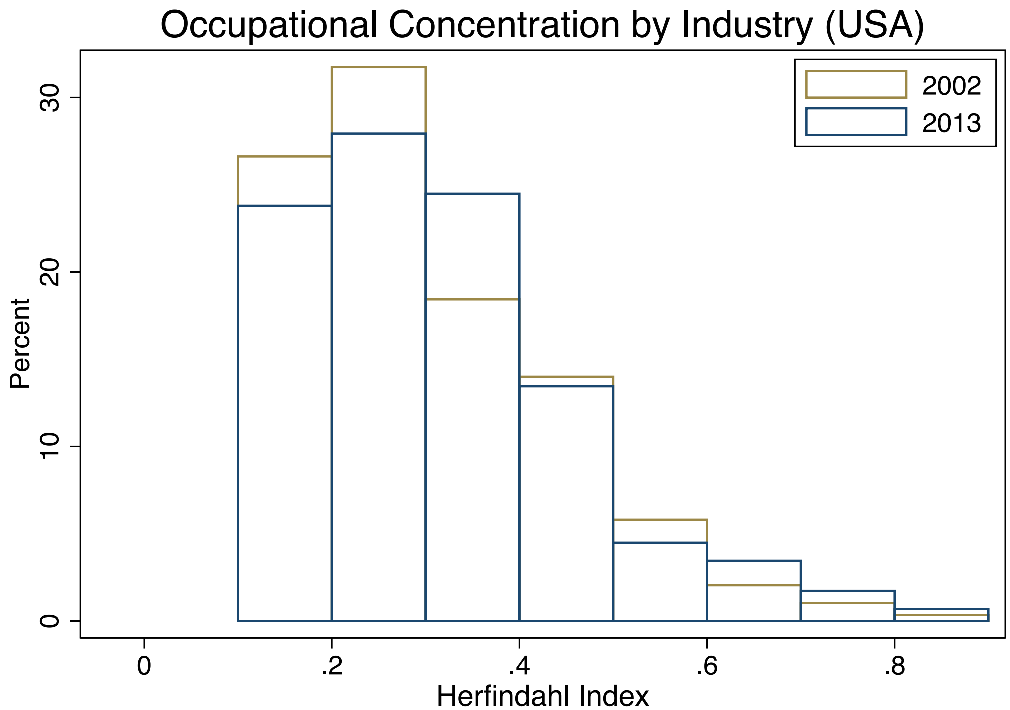

Figure 10

employment herfindahl across industries between 2002 and 2013.

2.4 Occupational Concentration and Wages

We also examine how changes in concentration affect occupation-level average wages in the OES. We find that low skill occupations tend to experience decreasing average wages in the concentration of those occupations, while high skill occupations tend to see increasing average wages as the concentration of those occupations increase. We also see definite patterns across occupations in terms of which ones tend to be outsourced compared to others.

2.4.1 Occupation-Level Wage Change and Outsourcing

The question of how outsourcing affects wages for different occupations, industries and skill levels has been studied at length over the last few decades. Holmes and Snider (2011) provide a theory of how decreases in an outsourcing friction may decrease wages for low-skill unionized labour, while Dube and Kaplan (2010) provide evidence that janitorial workers

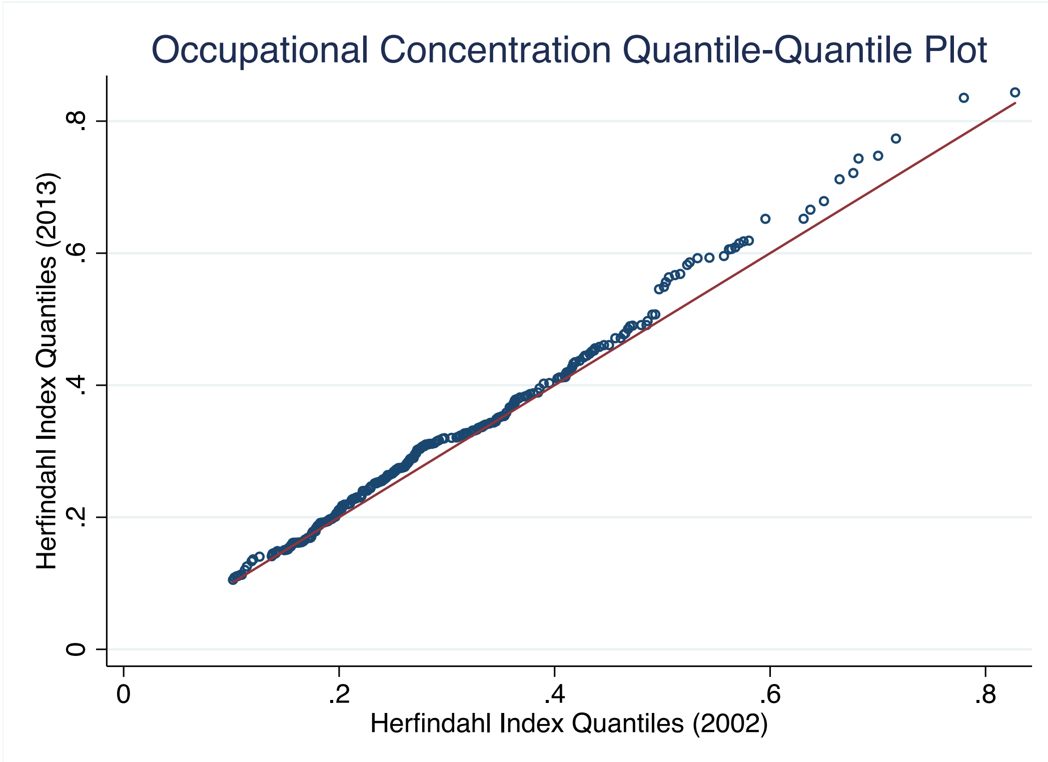

Figure 11

employed in the cleaning service industry receive lower wages/benefits than those employed in manufacturing. We find similar trends in the OES data, which we interpret as evidence for our argument that occupation sorting is an important component of changes in the wage distribution.

Table (1) summarizes occupation-level information on the relationship between occupation concentration within industries and the wage. The first two columns are the occupation code and group title. These represent fairly aggregate measures of occupation, but they can still be separated generally into higher skill/education occupations (roughly 11 through 29) and lower skill/education occupations (30+). The third column is a measure of whether or not average wages for an occupation in an industry are correlated with the total employment share for that occupation in that industry. A plus indicates a significant positive relationship, while a minus is a significant negative relationship. An empty space means that the relationship was not statistically different from zero. Here we see that occupation groups such as Management, Computer and Mathematical and Healthcare Practitioners tend to receive higher wages when they represent a higher proportion of employees within an industry, while Building and Grounds Cleaning, Protective Services and Office and Admin Support occupations tend to receive lower wages when they receive larger employment shares in an industry.

Table 1: Occupation Concentration and Wages in the OES

Code	**Occupation Group Title**	**Corr**	**Low Wage**	**% Increased employment share**	
				2002 to 2007	2002 to 2013
11-0000	Management	+	no	0.06	0.24
13-0000	Business and Financial Operations	+	no	**0.82**	**0.83**
15-0000	Computer and Mathematical	+	no	**0.50**	**0.62**
17-0000	Architecture and Engineering	+	**yes**	0.38	**0.51**
19-0000	Life, Physical, and Social Science		no	**0.57**	0.27
21-0000	Community and Social Service	-	**yes**	**0.54**	**0.56**
23-0000	Legal	-	no	**0.59**	**0.66**
25-0000	Education, Training, and Library		no	0.35	0.34
27-0000	Arts, Design, Entertainment, Sports, and Media	+	no	**0.59**	**0.57**
29-0000	Healthcare Practitioners and Technical	+	no	**0.54**	**0.71**
31-0000	Healthcare Support	-	**yes**	0.44	0.36
33-0000	Protective Service	-	**yes**	0.26	0.32
35-0000	Food Preparation and Serving Related		no	0.27	0.27
37-0000	Building and Grounds Cleaning and Maintenance	-	**yes**	0.19	0.11
39-0000	Personal Care and Service		no	0.43	0.44
41-0000	Sales and Related	-	**yes**	**0.60**	**0.58**
43-0000	Office and Administrative Support	-	no	0.44	0.29
45-0000	Farming, Fishing, and Forestry		**yes**	0.42	0.24
47-0000	Construction and Extraction	-	**yes**	0.29	0.24
49-0000	Installation, Maintenance, and Repair	+	no	0.47	0.46
51-0000	Production	-	**yes**	0.38	0.31
53-0000	Transportation and Material Moving	+	no	0.36	0.26

Corr - The correlation b/w occupation average industry wage and industry total employment share
Low Wage - Indicates whether or not the occupation-specific industries offer lower than average mean wages
The right most column shows percent of industries which increase employment share of that occupation

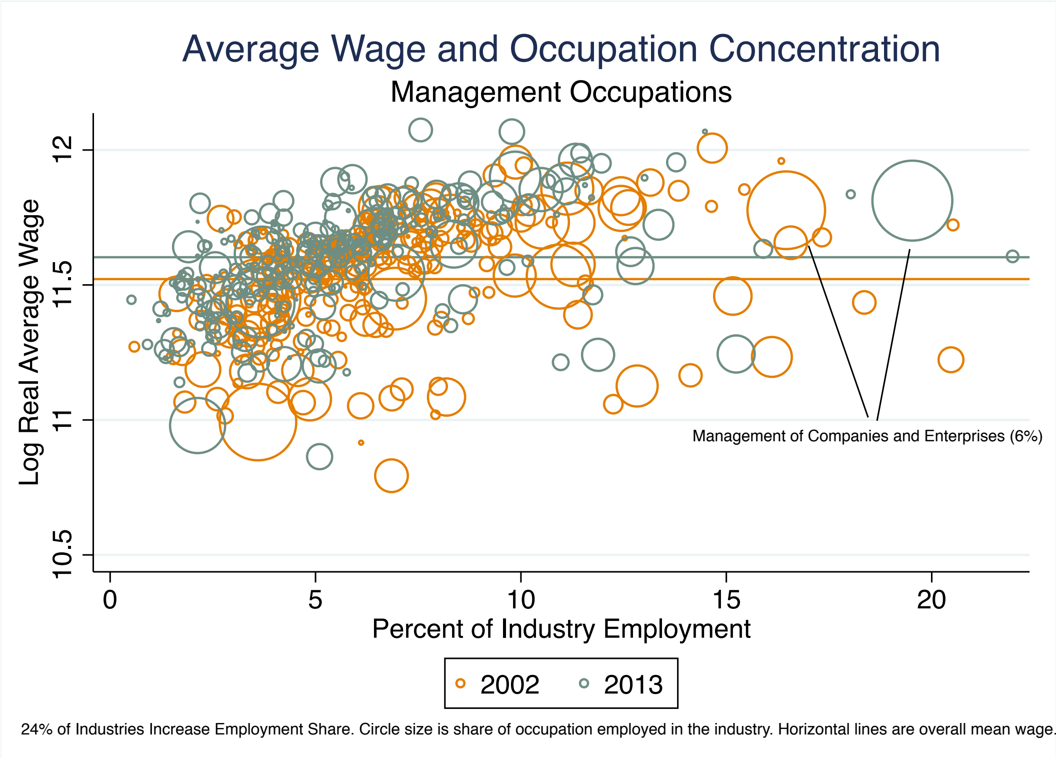

Figure 12

Most of these occupation groups can be linked to occupation-specific industries where the occupation produces the industry's primary output. For example, 37% of all workers in the Building and Grounds Cleaning and Maintenance occupation are employed by the Services to Buildings and Dwellings industry, and represent over 80% of all employment in that industry (see figure (14)). We identify the primary industries for each occupation group[15], and then check to see if the average wages received by the occupations in their primary industries are lower than the average occupation wage across all industries. The results are presented in Column 4 in table (1). Scientists working in the "Scientific Research and Development Services" industry make above average wages for scientists, while security guards (Protective Services) working in "Investigation and Security Services" make less than average wages for their occupation.

For each occupation, we also tabulate the proportion of industries with increase the share

[15]There's no exact way to identify such industry-occupation matches. Generally, for a given occupation, we identify which industries have large employment shares of that occupation, and in which the primary output matches the occupation. For example, the "Management of Companies and Enterprises" and "Land Subdivision" industries both consist of about 20% managerial employees, but we only include the former as a primary industry for management occupations.

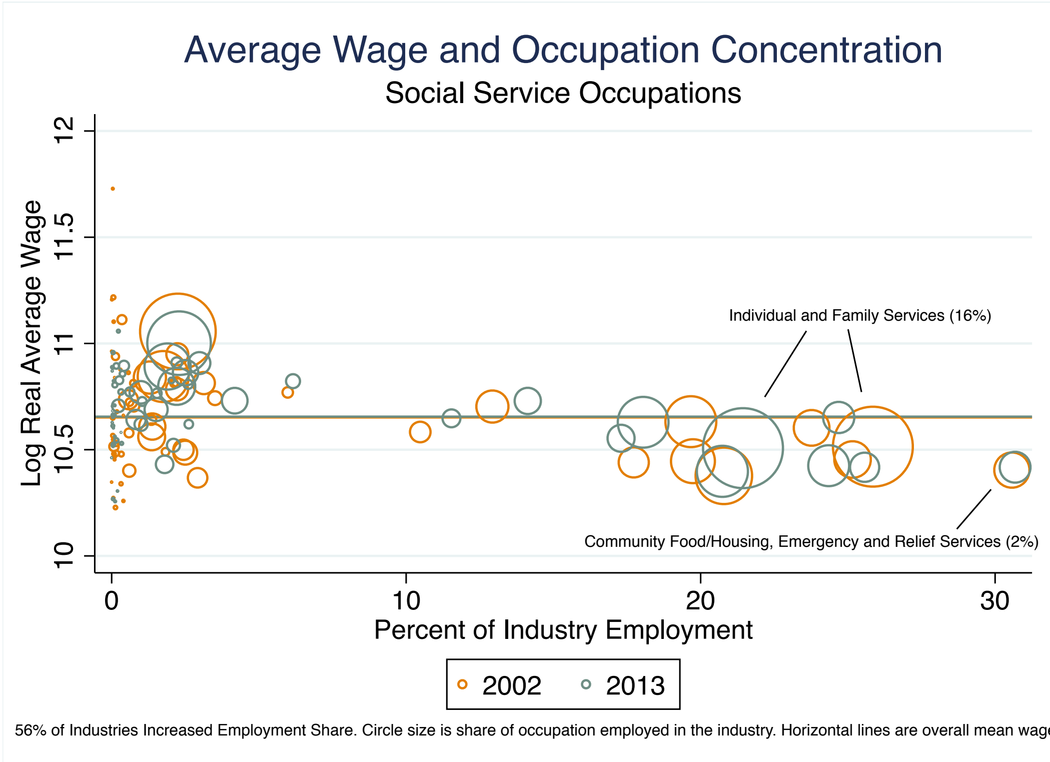

Figure 13

of employment going to that occupation, shown in columns 5 and 6. This can be seen as a measure of outsourcing. For example, 83% of all industries increased employment of Business and Financial Operations workers relative to other workers between 2002 and 2013, while only 11% of industries increased the employment share of Building and Grounds Cleaning and Maintenance workers.

Generally, the occupation groups can be split into two groups. It seems as if higher skill occupations are more likely to have wages increasing in concentration, while lower skill occupations seem more likely to decrease in concentration. As we discuss in the theoretical section, this could be due to differences in economies of scale or between-employee externalities generated by different occupations, or differences in labour demand elasticities across industries and occupations. Higher skill occupations also tend to receive higher than average wages in their specialized industries, while low skill occupations are the opposite. This fits the mechanism in our model, where primary occupation workers with the positive externalities receive higher wages as outsourcing frictions decrease, while workers without see wage declines in their primary industry. Finally, high skill workers seem to be capturing larger shares of employment across most industries relative to low skill occupations, which

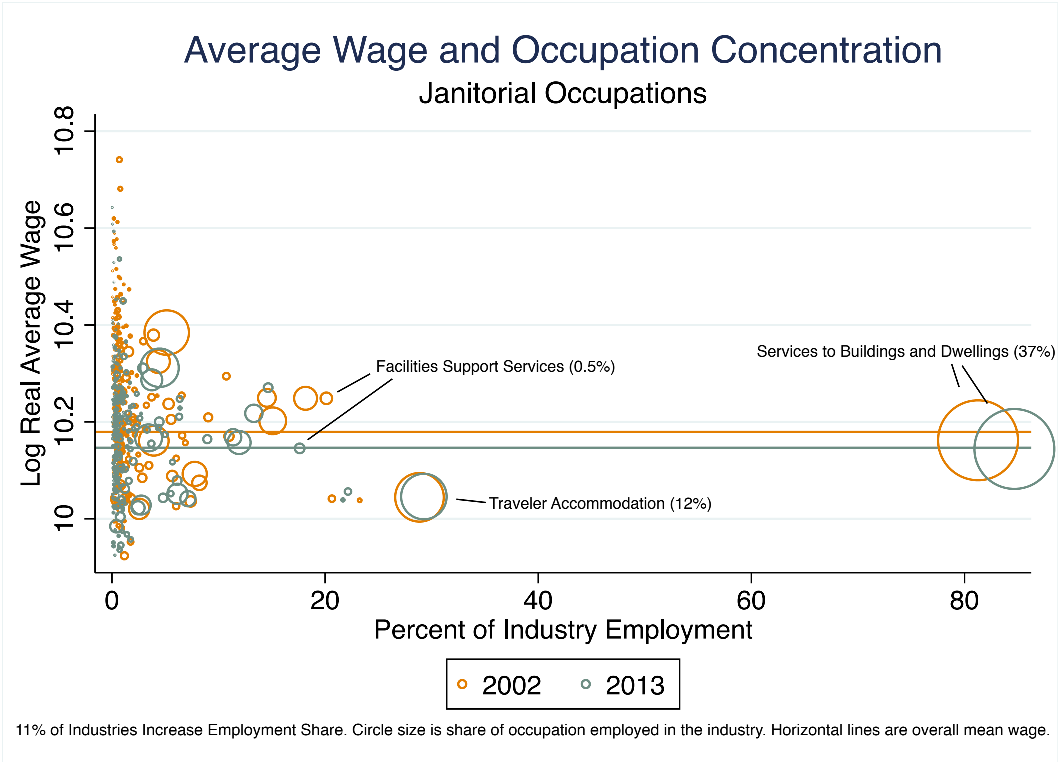

Figure 14

are growing in share in fewer industries.

These trends can be seen in more detail when we look at specific occupational groups. Figures 12, 13 and 14 show wage and employment trends for management, social service and janitorial occupations in 2002 and 2013. Each circle is an occupation-industry average wage observation, where the y-axis is log real average wage, and the x-axis is the employment share in that industry for that occupation. The size of the circle represents what proportion of the occupation is employed in that industry. The horizontal lines represent the average wages for that occupation in each year (management wages go up, social service wages stay constant, and Janitorial wages decline). Casual observation of the management and social service occupation plots verifies the positive and negative correlations, respectively, between average wage and industry employment share. Wages for managers in the management industry are clearly above average, and below average for social service workers in their primary industries.

2.5 Trends in Occupation Sorting into Industries

The previous sections investigate how the mix of occupations employed within each industry is changing over time, using the Herfindahl Index as a measure of concentration. We can look at this increase in occupational sorting from the other direction as well, by considering changes in the mix of industries in which particular occupations work. To do this, we calculate the Occupational Herfindahl Index, which is defined for occupation o as $H_o = \sum_{i \in I_o} s_{oi}^2$, where s_{oi} is the share of occupation o employed in industry i, and I_o is the set of industries in which occupation o is employed.

Table (2) summarizes the results for the 22 two-digit occupation groups. The groups are sorted by largest percentage change in Occupational Herfindahl. An increase in the Occupational Herfindahl represents an increase in sorting of that occupation across industries. Either the set of industries in which an occupation works has decreased, or fewer industries employ larger shares of that occupation. The results mirror the results with the Industry Herfindahl, in that about 2/3rds of occupations experience increases in industry segregation, with the weighted mean Occupational Herfindahl increasing from 0.112 to 0.128. The largest increase in segregation is experienced by the Farming, Fishing and Forestry occupation group, followed by the Science Occupation group. To put this in perspective, an increase from 0.054 to 0.076 is as if scientists were equally employed by 20 different industries, and then 7 industries outsourced all scientists to the remaining 13 industries (35% of the occupation).

Table 2: Changes in Occupation Concentration across Industries

Code	**Occupation Group Title**	**Labour Share**		**Occupational Herfindahl**		
		2002	2013	2002	2013	Change
45-0000	Farming, Fishing, and Forestry	0.3%	0.3%	0.239	0.353	47.5%
19-0000	Life, Physical, and Social Science	0.8%	0.9%	0.054	0.076	41.9%
15-0000	Computer and Mathematical	2.2%	2.8%	0.062	0.086	37.3%
37-0000	Building and Grounds Cleaning and Maintenance	3.3%	3.2%	0.132	0.169	27.9%
11-0000	Management	5.6%	4.9%	0.013	0.017	27.1%
35-0000	Food Preparation and Serving Related	7.9%	9.0%	0.438	0.513	17.2%
17-0000	Architecture and Engineering	1.9%	1.8%	0.080	0.092	15.4%
51-0000	Production	8.4%	6.6%	0.017	0.019	15.3%
39-0000	Personal Care and Service	2.3%	3.0%	0.057	0.065	14.5%
53-0000	Transportation and Material Moving	7.4%	6.8%	0.023	0.026	10.7%
41-0000	Sales and Related	10.5%	10.6%	0.029	0.031	8.1%
49-0000	Installation, Maintenance, and Repair	4.1%	3.9%	0.025	0.026	3.6%
23-0000	Legal	0.7%	0.8%	0.362	0.374	3.3%
13-0000	Business and Financial Operations	3.7%	5.0%	0.030	0.031	0.9%
33-0000	Protective Service	2.3%	2.5%	0.241	0.240	-0.4%
25-0000	Education, Training, and Library	6.1%	6.3%	0.476	0.465	-2.3%
21-0000	Community and Social Service	1.2%	1.4%	0.083	0.080	-3.5%
43-0000	Office and Administrative Support	17.9%	16.2%	0.018	0.017	-3.8%
27-0000	Arts, Design, Entertainment, Sports, and Media	1.2%	1.3%	0.036	0.035	-4.0%
31-0000	Healthcare Support	2.5%	3.0%	0.106	0.099	-7.1%
47-0000	Construction and Extraction	4.8%	3.8%	0.089	0.083	-7.1%
29-0000	Healthcare Practitioners and Technical	4.8%	5.9%	0.190	0.169	-10.9%
Total		100.0%	100.0%	0.112	0.128	14.3%

The Occupational Herfindahl Index is defined for occupation o as $H_o = \sum_{i \in I_o} s_{oi}^2$*, where* s_{oi} *is the share of occupation o employed in industry i, and* I_o *is the set of industries in which occupation o is employed.*

3 Theory

In order to better understand the relationships and trends in the data, we build a parsimonious model of inter-firm trade which explains how firms choose their inputs and the link between employment, wages and outsourcing. The model follows in the tradition of the heterogeneous-firm trade literature, with several key innovations: endogenous heterogeneous labour supply/demand and firm-level make or buy decisions. The basic environment is one with a network of multiple industries[16], each with a continuum of firms in monopolistic competition. Firms are differentiated by their unique good, and their occupation-specific labour productivities, where occupations are the primary labour type of their respective industry. Firms produce using their industry's primary occupation and a set of intermediates, which can be bought from other firms or made in-house by hiring labour of the appropriate occupation[17]. Trade in intermediates (goods and services) between firms incurs iceberg trade costs, representing outsourcing frictions. Occupations differ in productivity, while industries differ in their production technologies. Consumers have preferences over variety and also supply labour of each occupation type endogenously.

3.1 Notation

Writing down models of trade between industries and firms requires careful notation in order to avoid confusion for the readers (and the authors). We follow the conventions in the literature where possible. In general, subscripts refer to industries and sectors, while firms are indexed within brackets. For example, $P_i(f_i)$ is the price set by firm f in industry i (which we refer to as firm or good f_i). Double subscripts or bracketed indexes refer to trade. Typically, the first index refers to the destination, while the second refers to the source. τ_{ik} is the iceberg trade friction incurred by firms in industry i when they purchase intermediates from industry k. $q_{ik}(f_i, f_k)$ is the quantity demanded of good f_k by firm f_i.

[16]See De Soyres (2015) for a related network model of international trade.

[17]For example, a law office uses lawyers to produce its primary output (law services). The law office also requires cleaning services and furniture. It may either purchase these goods and services from other janitorial service and furniture manufacturing/retail firms, or it can hire in-house janitors and carpenters to provide them directly

3.2 Production

We consider an economy with a finite number N of industries, where each industry i has a unit mass of firms Ω_i. Each firm f in i produces its own differentiated variety of the industry-specific commodity/service $y_i(f)$ using a CRS production function in intermediates $M_i(f)$ and their primary labour type:

$$q_i(f_i) = z_i M_i(f_i)^{1-\beta_i} \ell_{ii}(f_i)^{\beta_i} \tag{5}$$

where z_i represents industry-specific efficiency and $\ell_{ii}(f_i)$ is quantity of occupation-type i used in industry i by firm f_i. $M_i(f_i)$ is a CES bundle of intermediate goods and services which are combined with industry-specific share parameters α_{ik}:

$$M_i(f_i) = \left(\sum_{k \neq i}^{N} \alpha_{ik}^{\frac{1}{\rho}} m_{ik}(f_i)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \tag{6}$$

where $\sum_{k \neq i} \alpha_{ik} = 1$, $\rho \geq 1$, and $m_{ik}(f_i)$ is the quantity of intermediate type k used by firm f_i.

Firms either purchase these intermediates from other firms for price P_{ik} or hire occupation-specific labour, with wage w_{ik}, to make it in-house[18]. Each firm has a vector of labour efficiency terms $\tilde{z}(f_i) = \{z_1(f_i), \ldots, z_N(f_i)\}$, so the unit cost to firm f_i for acquiring intermediate k is

$$C_{ik}(f_i) = \min \left\{ \tau_{ik} P_k, \frac{w_{ik}}{z_{ik}(f_i)} \right\} \tag{7}$$

where

$$m_{ik}(f_i) = \begin{cases} \tau_{ik}^{-1} q_{ik}(f_i) & \text{if } \tau_{ik} P_k \leq \frac{w_{ik}}{z_{ik}(f_i)} \\ z_{ik}(f_i) \ell_{ik}(f_i) & \text{if } \tau_{ik} P_k > \frac{w_{ik}}{z_{ik}(f_i)} \end{cases} \tag{8}$$

Here τ_{ik} represents a trade friction which results when purchasing intermediate k for use in industry i. We think of this friction as representing factors such as agency or contracting costs, transportation costs or customization/alteration costs. Such costs can be mitigated through technological progress which reduce the costs of outsourcing intermediate production relative to the cost of making it in-house. Here the friction is modeled similarly to an iceburg trade cost, where in order to get $m_{ik}(f_i)$ of usable intermediate k, the firm must purchase $q_{ik}(f_i) = \tau_{ik} m_{ik}(f_i)$ of the intermediate from industry k at a per-unit cost of $\tau_{ik} P_k$.

[18] We show later that the price for industry aggregates are the same for all firms. At this point, for simplicity, we assume the wage of k in i does not differ across firms.

When firms purchase intermediates, we assume they purchase some quantity of the CES industry aggregate, so

$$q_{ik}(f_i) = \left(\int_{\Omega_k} q_{ik}(f_i, f_k)^{\frac{\gamma_k - 1}{\gamma_k}} df_k \right)^{\frac{\gamma_k}{\gamma_k - 1}} \tag{9}$$

where $q_{ik}(f_i, f_k)$ is the quantity of good f_k purchased by f_i.

We assume that the labour productivities are drawn independently from a pareto distribution, with common shape parameter θ and occupation-specific lower bound $T_k^{-\theta}$. Thus, the probability ζ_{ik} that any given firm in industry i will hire labour to produce intermediate k is

$$\zeta_{ik} \equiv \Pr\left[z_{ik} > \frac{w_{ik}}{\tau_{ik} P_k} \right] = T_k \left(\frac{w_{ik}}{\tau_{ik} P_k} \right)^{-\theta} \tag{10}$$

Note that this cost/production structure implies that every firm will either make or buy any given intermediate - not both. However, each firm makes $N-1$ such decisions, and as such there are 2^{N-1} different firm configurations. In addition (and importantly), since each industry has a unit mass of firms, using a LLN we get that equation (10) is also the share of firms in i who will hire occupation k. This feature of the model will allow us to match industry-level data where we see both purchase of intermediate goods/services and the employment of the same type of intermediate labour within a given industry.

This simple structure already provides some useful intuition. First, as the price of inputs increases (decreases), the share of firms which outsource will decrease (increase). Alternately, outsourcing is increasing in the wage of the outsourced occupation. Firms which are vertically integrated will tend to be of higher productivity and will be larger, both in employment and output (due to lower costs and thus lower prices).

3.3 Prices and Trade

Firms incur a trade cost when buying intermediates. This cost, which we model as an iceberg cost, can represent physical transportation costs, as in the trade literature, as well as any variety of contracting, agency or efficiency costs associated with outsourcing a production process or service to an outside entity. Here, the delivery of one unit of an intermediate from a firm in k to a firm in i requires the shipment of $\tau_{ik} \geq 1$ units. Since firms are monopolistic competitors over their within-industry variety, they set prices at an industry-specific constant markup $\frac{\gamma_i}{\gamma_i - 1}$ over marginal cost $C_i(f_i)$. Thus, the price set by firm f_i is

$$P_i(f_i) = \bar{m}_i C_i(f_i) = \frac{\gamma_i}{\gamma_i - 1} \frac{P_i^B(f_i)}{z_i} \tag{11}$$

and the price faced for this good by any firm in industry k is

$$P_{ki}(f_i) = \tau_{ki} P_i(f_i) = \tau_{ki} \frac{\gamma_i}{\gamma_i - 1} \frac{P_i^B(f_i)}{z_i} \tag{12}$$

where $P_i^B(f_i)$ is the price of one unit of the optimal input bundle

$$P_i^B(f_i) = (1 - \beta_i)^{(\beta_i - 1)} \beta_i^{-\beta_i} P_i^I(f_i)^{1-\beta_i} w_{ii}^{\beta_i} \tag{13}$$

and $P_i^I(f_i)$ is the price of one unit of the optimal intermediate bundle:

$$P_i^I(f_i) = \left[\sum_{\ell \neq i}^{N} \alpha_{i\ell} C_{i\ell}(f_i)^{1-\rho} \right]^{\frac{1}{1-\rho}} \tag{14}$$

The price of good i in industry k is then the CES price index[19]

$$P_i = \left(\int_{\Omega_i} P_i(f_i)^{1-\gamma_i} df_i \right)^{\frac{1}{1-\gamma_i}} \tag{15}$$

Note that these price indices are in terms of the optimal bundle at the given prices. Since we are integrating out across f in i, the price of the intermediate industry aggregate is the same for all firms in industry k. For ease of analysis, we assume for now that $\rho = 1$, so the intermediate aggregate and price index become:

$$M_i(f_i) = \prod_{\ell \neq i} m_{i\ell}(f_i)^{\alpha_{i\ell}} \tag{16}$$

and

$$P_i^I(f_i) = \prod_{\ell \neq i} \left(\frac{C_{i\ell}(f_i)}{\alpha_{i\ell}} \right)^{\alpha_{i\ell}} \tag{17}$$

Since firms differ solely on their variety and their vector of labour productivity terms $\tilde{z}(f)$, we can express the cost of inputs as: $C_{ik}(\tilde{z}) = \min\{\tau_{ik} P_k, w_{ik}/z_k\}$, which we then plug

[19] $P_i^B(f_i)$ is derived from the cost-minimization problem of the firm over the intermediate index and primary labour. $P_i^I(f_i)$ is derived from the cost-minimization problem over intermediates.

into equations 14, 13 and 12 to get

$$P_i(\tilde{z}) = \psi_i w_{ii}^{\beta}\bigg(\prod_{\ell}^{N} \alpha_{i\ell}^{-\alpha_{i\ell}} \min\{\frac{w_{i\ell}}{z_\ell}, \tau_{i\ell}P_\ell\}^{\alpha_{i\ell}}\bigg)^{1-\beta_i} \tag{18}$$

where $\psi_i = \frac{\gamma_i}{\gamma_i - 1}(1-\beta_i)^{\beta_i - 1}\beta_i^{-\beta_i}z_i^{-1}$. So, the price of a firm's good within an industry depends on the industry's technology parameters $\alpha_{i\ell}$, the wages of all the different occupations within that industry $w_{i\ell}$, industry productivity, firm labour productivity and the price of the industry aggregate for all the other industries, which in turn depend on all the firm prices within those industries. This price only varies across recipient industries through the trade cost parameter τ_{ki}.

3.4 Labour Demand and Intermediates

Firms which buy intermediates combine within-industry varieties according to the CES aggregator in equation (9), implying that they buy a little bit from every firm in that industry. Demand for good j in k by firm f in i is

$$q_{ik}(f_i, f_k) = \frac{P_k(f_k)^{-\gamma_k}}{P_k^{1-\gamma_k}} \int_{\Omega_k} P_k(j_k)q_{ik}(f_i, j_k)dj_k = \frac{P_k(f_k)^{-\gamma_k}}{P_k^{1-\gamma_k}} X_{ik}^P(f_i) \tag{19}$$

where $X_{ik}^P(f_i)$ is total expenditure on purchasing intermediate k by f_i. Note that if $z_k(f_i) > w_{ik}/(\tau ikP_k)$, then $X_{ik}^P(f_i) = 0$. Define $\Omega_{ik}^P = \{f_i \in \Omega_i | z_k(f_i) \leq w_{ik}/(\tau_{ik}P_k)\}$ as the measure of firms in i which purchase intermediate k. The demand for good f_k from industry i is

$$q_{ik}(f_k) = \int_{\Omega_{ik}^P} q_{ik}(f_i, f_k)df_i = \int_{\Omega_{ik}^P} \frac{P_k(f_k)^{-\gamma_k}}{P_k^{1-\gamma_k}} X_{ik}^P(f_i)df_i = \frac{P_k(f_k)^{-\gamma_k}}{P_k^{1-\gamma_k}} X_{ik}^P \tag{20}$$

and total demand for good $k(j)$ is

$$q_k(f_k) = q_{ck}(f_k) + \sum_{i \neq k} q_{ik}(f_k) = q_{ck}(f_k) + \sum_{i \neq k} \frac{P_k(f_k)^{-\gamma_k}}{P_k^{1-\gamma_k}} X_{ik}^P \tag{21}$$

where $q_{ck}(f_k)$ is consumption demand for good f_k[20]. X_{ik}^P is total expenditure on intermediate k by those firms in i which outsource (buy) k.

In order to proceed further, we need to define firm-level demand for aggregate interme-

[20] We derive consumption demand in the next section

diates and labour in terms of demand for its own output. To do this, we first use the firm's cost-minimizing problem (which is to minimize $P_i^I(f_i)M_i(f_i) + w_{ii}\ell_{ii}(f_i)$ subject to (5)) to get

$$M_i(f_i) = \left(\frac{1-\beta_i}{\beta_i}\right)^{\beta}_{i} \left(\frac{w_{ii}}{P_i^I(f_i)}\right)^{\beta}_{i} z_i^{-1} q_i(f_i) \tag{22}$$

$$\ell_{ii}(f_i) = \left(\frac{1-\beta_i}{\beta_i}\right)^{\beta_i - 1} \left(\frac{w_{ii}}{P_i^I(f_i)}\right)^{\beta_i - 1} z_i^{-1} q_i(f_i) \tag{23}$$

We can then get demand for intermediates by minimizing $\sum_{k\neq i}^{N} C_{ik}(f_i)m_{ik}(f_i)$ subject to (16), which results in

$$\begin{aligned}
m_{ik}(f_i) &= \frac{\alpha_{ik}}{C_{ik}(f_i)} \prod_{\ell\neq k}^{N} \left(\frac{C_{i\ell}(f_i)}{\alpha_{i\ell}}\right)^{\alpha_{i\ell}} M_i(f_i) && (24)\\
&= \frac{\alpha_{ik}}{C_{ik}(f_i)} P_i^I(f_i) M_i(f_i) && (25)\\
&= \frac{\alpha_{ik}}{C_{ik}(f_i)} P_i^I(f_i) \left(\frac{1-\beta_i}{\beta_i}\right)^{\beta_i} \left(\frac{w_{ii}}{P_i^I(f_i)}\right)^{\beta_i} z_i^{-1} q_i(f_i) && (26)\\
&= \frac{\alpha_{ik}}{C_{ik}(f_i)} (1-\beta_i) P_i^B(f_i) z_i^{-1} q_i(f_i) && (27)
\end{aligned}$$

Note that this gives us the share of intermediate k in $M_i(f_i)$

$$\frac{m_{ik}(f_i)}{M_i(f_i)} = \frac{\alpha_{ik}}{C_{ik}(f_i)} P_i^I(f_i) \tag{28}$$

and allows us to get an expression for labour and overall intermediate demand, with

$$\ell_{ik}(f_i) = \begin{cases} \frac{m_{ik}(f_i)}{z_k(f_i)} = \frac{\alpha_{ik}}{w_{ik}}(1-\beta_i)P_i^B(f_i) z_i^{-1} q_i(f_i) & \text{if } z_k(f_i) > \frac{w_{ik}}{\tau_{ik}P_k} \\ 0 & \text{if } z_k(f_i) \leq \frac{w_{ik}}{\tau_{ik}P_k} \end{cases} \tag{29}$$

and

$$m_{ik}(f_i) = \begin{cases} \alpha_{ik}\left(\frac{w_{ik}}{z_k(f_i)}\right)^{-1}(1-\beta_i)P_i^B(f_i) z_i^{-1} q_i(f_i) & \text{if } z_k(f_i) > \frac{w_{ik}}{\tau_{ik}P_k} \\ \alpha_{ik}(\tau_{ik}P_k)^{-1}(1-\beta_i)P_i^B(f_i) z_i^{-1} q_i(f_i) & \text{if } z_k(f_i) \leq \frac{w_{ik}}{\tau_{ik}P_k} \end{cases} \tag{30}$$

So, given $q_i(f_i)$, wages, prices and labour productivity, the firm will choose its ideal mix of hiring and outsourcing according to (23), (29) and (30).

3.5 Consumption and Labour Supply

Consumers in our model combine goods using a CES aggregator as is standard in much of the trade literature such as, recently, Eaton, Kortum and Kramarz (2011). In order to close the model so that we can do counterfactuals, we also need to specify household preferences over labour and occupations. Most of the trade literature sidesteps this issue with a fixed supply of homogeneous labour within in each country which is perfectly mobile across firms but not across borders. In a setting with multiple occupations and general preferences over labour, the representative agent will substitute across labour markets such that wages equalize. Since our research is specifically interested in wages differences across industries and occupations, we need our model to be able to replicate the observed differences in average occupation wages across firms and industries. However, we wish to do this in as parsimonious a fashion as possible[21]. To this end, we model our household as providing labour across occupations and industries using a CES aggregator similar to the consumption aggregator but with labour supply elasticity parameter ν. The interpretation for these preferences is that labour is not perfectly mobile across occupations, industries and job locations due to differences in supply, training, preferences, location, ability and other personal constraints (represented in somewhat reduced form by ϕ_{ik}, where $\sum_{ik} \phi_{ik} = 1$). Thus, while labour supply is guided by relative wages, wages will not necessarily equate in equilibrium. In fact, this simple specification is enough to generate exactly the kind of wage dynamics we see in the data.

Households own the firms, and demand consumption goods and supply labour in order to maximize

$$\prod_{i}^{N} \left(\int_{\Omega_i} q_{ci}(f_i)^{\frac{\sigma_i - 1}{\sigma_i}} \right)^{\frac{\sigma_i}{\sigma_i - 1}\mu_i} - \psi \left(\sum_{i,k} \phi_{ik}^{\frac{1}{\nu}} \ell_{ik}^{\frac{1+\nu}{\nu}} \right)^{\frac{\nu}{1+\nu}} \tag{31}$$

subject to

$$\sum_{i}^{N} \int_{\Omega_i} P_{ci}(f_i) q_{ci}(f_i) = \sum_{i,k} w_{ik}\ell_{ik} + \Pi \tag{32}$$

where Π are total firm profits (which are positive due to monopolistic competition), ψ is a level shifter, and $P_{ci}(f_i)$ is the price of good f_i for final consumption. μ_i is a consumption preference parameter over industry-types, and $\sum_i^N \mu_i = 1$. Note that in general we allow $P_{ci}(f_i)$ to differ from intermediate prices since the consumption elasticity parameter σ may differ from the intermediate elasticity γ, and trade friction τ_{ci} may be greater than one.

[21] We have also examined this question using a related 2x2x2 model where wage setting is done via union bargaining, as in Holmes and Snider (2011). See the appendix for details and results from that model.

Solving the household's problem provides consumption demand

$$q_{ci}(f_i) = \frac{P_i(f_i)^{-\sigma_i}}{P_i^{1-\sigma_i}}\mu_i P_c q_c = \frac{P_i(f_i)^{-\sigma_i}}{P_i^{1-\sigma_i}}\mu_i(WL + \Pi) \tag{33}$$

where $WL = \sum_{i,k} w_{ik}\ell_{ik}$ is total labour income, P_c is the consumption price index, defined similarly to the intermediate price index, q_c is the consumption aggregator, and consumption of good f_i is a fraction of total consumption expenditure $(WL + \Pi) = X_c$. Note that $\mu_i(WL+\Pi) = X_{ci}$ is the fraction of total consumption expenditure used on industry-type i.

Similarly, we can solve for household labour supply, which results in

$$\ell_{ik}^s = \frac{w_{ik}^\nu}{W^{1+\nu}}\phi_{ik} WL \tag{34}$$

where the wage aggregator W is defined as $W = (\sum_{i,k}\phi_{ik}^{-1} w_{ik}^{1+\nu})^{\frac{1}{1+\nu}}$. This specification gives an elasticity of labour supply equal to $(1 + \nu) > 0$, so labour supply is increasing in own wage.

3.6 Industry Aggregation

Though this model is suitable for firm-level analysis, such as in Eaton, Kortum and Kramarz (2011), we are primarily interested in industry-occupation level dynamics and trends. Fortunately, our model provides a parsimonious aggregation from firm to industry, despite the complex nature of firm outsourcing and hiring behaviour.

In most models of international trade[22], firms heterogeneously make entry/export decisions based on a single efficiency or productivity parameter. Since firms in our model draw $N - 1$ individual efficiency parameters, integrating out across firms may seem more difficult. In practice, since each draw is independent, we can still easily solve for industry-level quantities.

In order to do the aggregation, first note that as in equation (18) we can express firm-level prices in terms of the vector of productivity parameters $\tilde{z}$. Thus we can write the

[22]See Eaton and Kortum (2002), Chaney (2008)

industry-level price index as

$$P_i = \left(\int_{\Omega_{\tilde{z}}} P_i(\tilde{z})^{1-\gamma_i} dF(\tilde{z}) \right)^{\frac{1}{1-\gamma_i}} \tag{35}$$

$$= \left[\int_{\Omega_{\tilde{z}}} \left(\psi_i w_{ii}^{\beta_i} (\prod_{\ell}^{N} \alpha_{i\ell}^{-\alpha_{i\ell}} \min\{\frac{w_{i\ell}}{z_\ell}, \tau_{i\ell} P_\ell\}^{\alpha_{i\ell}})^{1-\beta_i} \right)^{1-\gamma_i} dF(\tilde{z}) \right]^{\frac{1}{1-\gamma_i}} \tag{36}$$

Using the fact that each of z_ℓ terms are drawn independently, we can integrate across all $2^{(N-1)}$ combinations of firm configurations to get the following N system of equations in industry prices:

$$P_i = \psi_i w_{ii}^{\beta_i} (\prod_{\ell \neq i} \alpha_{i\ell}^{-\tilde{\alpha}_{i\ell}}) \left[\prod_{\ell \neq i} \left(1 - \frac{\tilde{\alpha}_{i\ell}}{\theta + \tilde{\alpha}_{i\ell}} T_\ell (\frac{w_{i\ell}}{\tau_{i\ell} P_\ell})^{-\theta} \right) (\tau_{i\ell} P_\ell)^{\tilde{\alpha}_{i\ell}} \right]^{\frac{1}{1-\gamma_i}} \tag{37}$$

where $\tilde{\alpha}_{i\ell} = \alpha_{i\ell}(1-\beta_i)(1-\gamma_i)$ and consumer prices are obtained by using $\tau_{c\ell}$ and replacing γ_i with σ_i.

Similarly, we can integrate across firm labour demand (23) to get aggregate demand for labour type i in industry k:

$$\ell_{ki} = \int_{\Omega_z} \frac{\alpha_{ki}}{w_{ki}} (1-\beta_k) P_k^B(\tilde{z}) z_k^{-1} q_k(\tilde{z}) \mathbb{1}\{\tilde{z}_i > w_{ki}/(\tau_{ki} P_i)\} dF(\tilde{z})$$

$$= \int_{\Omega_z} \frac{\alpha_{ki}}{w_{ki}} P_k^I(\tilde{z})^{1-\beta_k} \left(\frac{1-\beta_k}{\beta_k} \right)^{\beta_k} w_{kk}^{\beta_k} z_k^{-1} \left(\tau_{ck} q_{ck}(\tilde{z}) + \sum_{\ell \neq k} \tau_{\ell k} q_{\ell k}(\tilde{z}) \right) \mathbb{1}\{\tilde{z}_i > w_{ki}/(\tau_{ki} P_i)\} dF(\tilde{z})$$

$$= \int_{\Omega_z} \frac{\alpha_{ki}}{w_{ki}} P_k^I(\tilde{z})^{1-\beta_k} \left(\frac{1-\beta_k}{\beta_k} \right)^{\beta_k} w_{kk}^{\beta_k} z_i^{-1} \frac{P_k(\tilde{z})^{-\gamma_k}}{P_k^{1-\gamma_k}} \left(X_{ck} + \sum_{\ell \neq k} X_{\ell k}^P \right) \mathbb{1}\{\tilde{z}_i > w_{ki}/(\tau_{ki} P_i)\} dF(\tilde{z})$$

$$= \frac{\alpha_{ki}(1-\beta_k)}{w_{ki}} \frac{\gamma_k - 1}{\gamma_k} \lambda_{ki} R_k \tag{38}$$

where

$$\lambda_{ki} \equiv \frac{w_{ki} \ell_{ki}}{X_{ki}} = \frac{\theta \zeta_{ki}}{\theta + \tilde{\alpha}_{ki}(1-\zeta_{ki})} \tag{39}$$

is the fraction of industry k expenditure on intermediate i which is spent on labour, and ζ_{ki} is the probability that a firm in k makes input i in-house as defined in equation 10. Specifically, $w_{ki}\ell_{ki} = \lambda_{ki} X_{ki}$. $R_k = X_{ck} + \sum_{\ell \neq k} X_{\ell k}^P$ is total revenues in industry k. Note that for simplicity, here we are assuming that $\sigma_k = \gamma_k$.

Notice that an increase in wages will have an affect on labour demand both on the intensive and extensive margins. As wages go up, the firms wish to keep their share of expenditure on that intermediate constant, and so they reduce amount of labour demanded. At the same time, increasing wages increases the productivity cutoff for vertical integration. Thus, increased wages decreases the measure of firms which make their own intermediate, with the marginal firms firing their labour and buying their intermediates on the open market instead.

Intuitively, as trade costs decrease, the productivity cutoff increases, and so labour demand for a particular occupation within an industry is increasing in trade costs (via input prices). Thus as trade costs decrease, demand for intermediate labour also decreases (holding wages and output demand fixed). Since we are able to solve for the equilibrium in this model, we will be able to also analyze how trade costs influence wages and demand in equilibrium as well.

Naturally, this result also gives us aggregate demand for purchased intermediates in industry k, which is

$$q_{ki} = \frac{\alpha_{ki}(1-\beta_k)}{\tau_{ki}P_i}\frac{\gamma_k - 1}{\gamma_k}(1-\lambda_{ki})R_k = \frac{(1-\lambda_{ki})X_{ki}}{\tau_{ki}P_i} = \frac{X_{ki}^P}{\tau_{ki}P_i} \tag{40}$$

Equation (40) allows us to solve for industry revenues R_k as follows:

$$R_k = X_{ck} + \sum_{i \neq k} X_{ik}^P \tag{41}$$

$$= \mu_k(WL + \Pi) + \sum_{i \neq k} P_k q_{ik} \tag{42}$$

$$= \mu_k(WL + \Pi) + \sum_{i \neq k} \alpha_{ik}(1-\beta_i)\frac{\gamma_i - 1}{\gamma_i}(1-\lambda_{ik})R_i \tag{43}$$

Since $\Pi_k = R_k/\gamma_k$, we can rewrite the expression for R_k as

$$R_k = \mathbb{A}_k WL + \sum_{i \neq k} \mathbb{B}_{ik} R_i \tag{44}$$

which is a linear system of equations in revenues, where $\mathbb{A}_k = (\mu_k\gamma_k)/(\gamma_k - \mu_k)$ and $\mathbb{B}_{ik} = (\mu_k + (1-\beta_i)(\gamma_i - 1)\alpha_{ik}(1-\lambda_{ik}))(\gamma_i - \mu_k)^{-1}$.

3.7 Comparative Statics in Partial Equilibrium

In order to illustrate the main dynamics at work in the model, we perform a few comparative statics exercises[23]. The main relationship of interest is that between changes in trade frictions τ_{ik} and employment dynamics. If we hold prices and wages constant in partial equilibrium, the main dynamics act through R_i (demand) and λ_{ik} (substitution). Given the assumptions of the model, holding prices fixed, we can show that $\partial\lambda_{ik}/\partial\tau_{ik} > 0$ and $\partial\lambda_{ik}/\partial\tau_{\ell j} = 0 \,\forall\, \ell, j \neq i, k$. Thus, decreases in directional trade frictions between industries i and k will result in increased outsourcing from i to k, while changes in trade costs between separate industries ℓ and j have no direct effect on trade and outsourcing between i and k. Of course, in the full general equilibrium exercise this will not be the case, as changes in any of the trade frictions will have network effects on outsourcing, wages and output in all of the other industries (depending on network linkages).

Similarly, we can show that $\partial R_k/\partial\tau_{ik} < 0$ and $\partial R_k/\partial\tau_{i\ell} < 0 \,\forall\, \ell, i, k$. The former is the *direct effect* of trade frictions, while the latter is the *feedback effect*. Both are negative, implying that decreasing trade costs increases revenues in all industries, holding wages and prices fixed. This is because trade frictions represent a dead-weight loss which constrains gains from trade and absolute/comparative advantage.

The employment effects of trade frictions in partial equilibrium are a little more complicated. Recall that

$$\ell_{ik} = \frac{\alpha_{ik}(1-\beta_i)}{w_{ik}}\frac{\gamma_i - 1}{\gamma_i}\lambda_{ik}R_i \quad \text{and} \quad \ell_{ii} = \frac{\beta_i}{w_{ii}}\frac{\gamma_i - 1}{\gamma_i}R_i$$

so a decrease in trade frictions from i to k will tend to increase primary-type employment in industry i,

$$\frac{\partial\ell_{ii}}{\partial\tau_{ik}} = \frac{\beta_i}{w_{ii}}\frac{\gamma_i - 1}{\gamma_i}\frac{\partial R_i}{\partial\tau_{ik}} < 0$$

However the effect on ℓ_{ik} is ambiguous, since the demand and substitution effects move in different effects. To fully estimate the effects of trade frictions on employment, wages, wage inequality and occupation sorting/concentration, we need to go to the general equilibrium framework of the full model.

[23] See appendix D for derivations and proofs.

3.8 Equilibrium and Analysis

Solving the model for all endogenous variables (prices, wages, quantities and labour) is then just a matter of solving the system of equations resulting from (34), (37), (38) and (44) along with a normalization of total labour supply.

The key relationship of interest in the model is between trade costs and wages/employment. It's easy to show that as general trade costs decrease, the share of firms in every industry which outsource increases. This reduces demand for intermediate labour in those industries, and increases demand for primary labour in every industry. If some occupations are paid more in their primary industry while others are paid less than average (as we show is the case in the data later in the paper), this will lead to increased wage variance and an increase in occupational concentration/sorting. As mentioned before, this model also nicely replicates several key facts about firm size and productivity, where more productive firms hire more labour, are less specialized and produce a larger share of output. Another nice feature of our model is that since it represents a network of interconnected firms across multiple industries, a shock to any individual firm or industry will echo through the network with an intensity related to the set of linkages. For example, if the mining industry experiences a negative productivity shock, this will propagate through the economy via prices (through equation 37), raising prices, lowering make/buy productivity cutoffs, and increasing labour demand even in industries which do not directly trade with the mining industry. Note also that the asymmetric nature of our trade frictions will affect the propagation of any shocks through the network.

The next step of our analysis is to use the model to estimate trade costs and productivity. In order to accomplish this, we use the industry-occupation level panel of wages, labour, intermediate inputs and output which we glean from industry input-output tables and the occupation-employment statistics discussed in the empirical section. Given these estimates, we then close the model and run counterfactuals to see how equilibrium wages and occupation allocations respond to changes in trade costs or productivity shocks, both at the sectoral or aggregate level.

3.9 Empirical Implications

The analysis of the model provides several predictions which we can compare to the empirical results in the data section. Specifically, decreases in common outsourcing costs due to technological or policy change should be associated with increasing Employment Herfindahl indexes within firms, industries and across the economy as a whole. Additionally, we should see increasing between-industry wage variance and declining within wage variance due to compositional change/sorting, and increases in within industry wage variance due to changes in wages. We could also potentially test the prediction that more productive firms are less specialized, or that changes in outsourcing are associated with overall decreases in labor demand or increases in aggregate output.

4 Model Calibration and Counter-factual Exercises

In this section, we discuss how we solve and calibrate the model using publicly available data on aggregate employment, wages and output at the industry-occupation level in the USA. We are able to use the model to back out trends in unobserved trade frictions, productivity and prices, which then allow us to perform several counter-factual exercises. The primary exercise is a decomposition where we hold trade frictions at their 2002 level in order to determine the effect of decreasing frictions on inequality and occupational concentration/sorting.

4.1 Data and Calibration

One convenient property of our model is that it's a firm-occupation level model which can easily be aggregated up to the industry level while retaining the firm-level outsourcing and trade dynamics. This allows us to take the model to industry-occupation level data and estimate the parameters which underly firm behaviour in the model and the aggregate trends we see in the data. In particular, we use industry and occupation level data from the OES (as discussed in the empirical section) and Input-Output make and buy tables from the BLS. Note that since this is a static model, we calibrate the model to the data period-by-period.

The first task is to create a mapping between industries and occupation groups. Our model implies that each industry has a particular set of occupations or labor types which produce the primary output of that industry. In order to map our model to the data, we define an aggregation and one-to-one mapping of occupations and industries which leaves us with $N = 20$ industry-occupation pairs. We then aggregate both the OES and the Input-Output tables to our pre-defined set of industries[24]

[24] See appendix A for details on how we construct the mapping and aggregate/construct the IO tables used

We then get (mean) wages and employment for occupation k in industry i, w_{ik} and ℓ_{ik}, from the OES. Similarly, we can read industry revenues (R_i) and purchased expenditures (X_{ik}^P, X_{ck}) directly from the IO tables. This allows us to construct total input expenditure $X_{ik} \equiv w_{ik}\ell_{ik} + X_{ik}^P$. Given wages, employment and input expenditure, we can use 39 to back out expenditure shares λ_{ik}. Accounting profits are constructed as revenues less input expenditures (we abstract away from investment or other firm expenditures in the model), which then gives us estimates of demand elasticities $\gamma_i = R_i/\Pi_i$. Given elasticities, we can use equation 38 to back out the scale parameters $\alpha_i k$ and β_i. This allows us to then back out outsourcing probabilities ζ_{ik} by inverting equation 39.

The labor-supply side parameters are recovered similarly. Given consumption expenditures for each industry's output, we can construct the demand terms μ_i as the consumption expenditure share for industry i. Solving for the labor supply/occupation preference terms ϕ_{ik} involves solving the linear system of N^2 equations represented by 34. Given the restrictions that the preference terms are positive and normalized to sum to one, we get a unique solution to the following system:

$$\begin{bmatrix} w_{11}^{1+\nu} - \frac{w_{11}^{\nu}}{\ell_{11}} WL & w_{12}^{1+\nu} & \cdots & w_{NN}^{1+\nu} \\ w_{11}^{1+\nu} & w_{12}^{1+\nu} - \frac{w_{12}^{\nu}}{\ell_{12}} WL & \cdots & w_{NN}^{1+\nu} \\ \vdots & & \ddots & \vdots \\ w_{11}^{1+\nu} & w_{12}^{1+\nu} & \cdots & w_{NN}^{1+\nu} - \frac{w_{NN}^{\nu}}{\ell_{NN}} WL \end{bmatrix} \begin{bmatrix} \frac{1}{\phi_{11}} \\ \frac{1}{\phi_{12}} \\ \vdots \\ \frac{1}{\phi_{NN}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{45}$$

This just leaves prices P_i, trade frictions τ_{ik}, and productivity terms z_i and T_i (which for now we normalize to 1). We solve these jointly using the equilibrium price system (37), the equation for ζ_{ik} (10) and the observation that

$$\frac{\tau_{ik}}{\tau_{jk}} = \frac{\zeta_{ik}^{-1/\theta} w_{ik}}{\zeta_{jk}^{-1/\theta} w_{jk}} \tag{46}$$

Given the calibrated model, we can then examine trends in trade frictions, concentration and wage inequality through the lens of the model.

in our calibration.

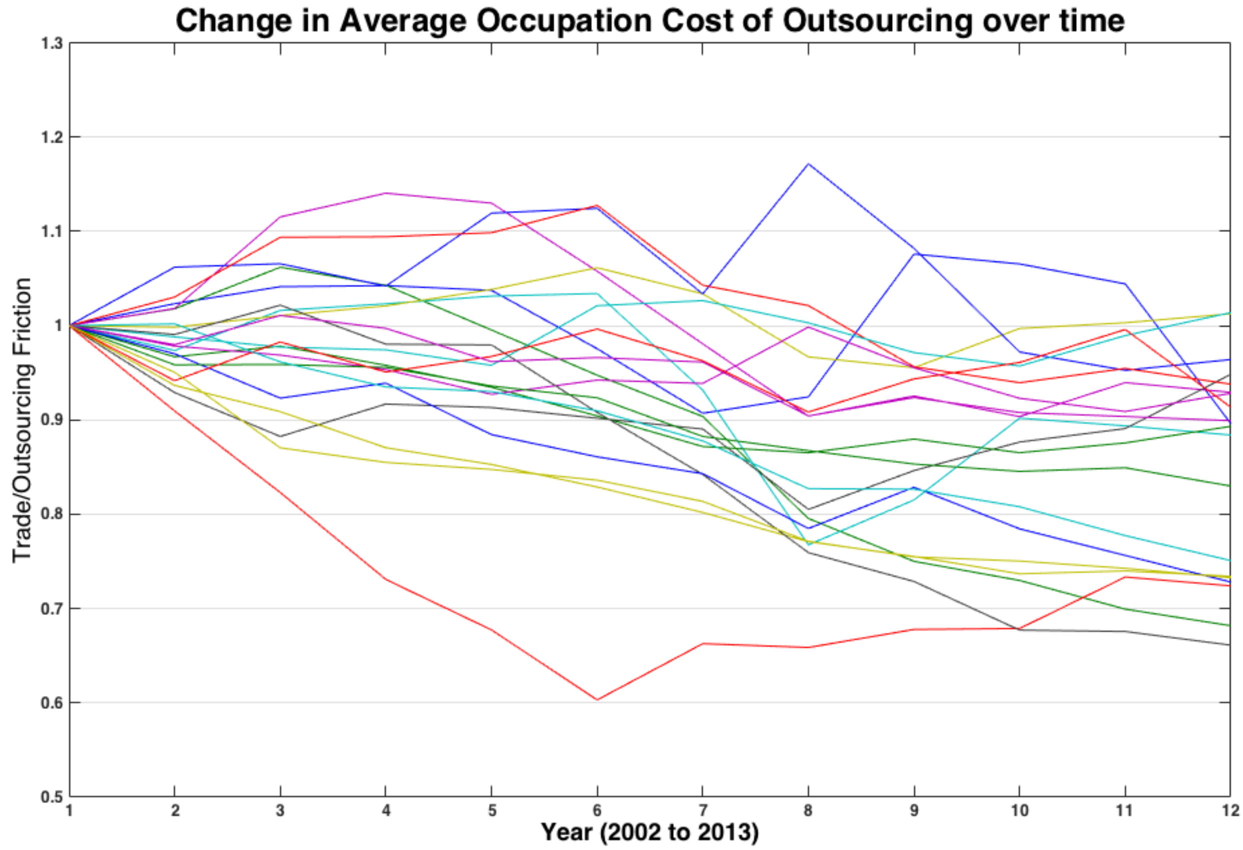

Figure 15: Change in the Average Occupation Cost of Outsourcing τ_k^O over time

4.2 Results

The first result from the model is that trade frictions declined significantly between 2002 and 2013. To illustrate these changes, we define the *Average Occupation Cost of Outsourcing* as the mean friction faced by all industries when attempting to outsource a particular input k. That is,

$$\tau_k^O = \frac{1}{N} \sum_i \tau_{ik} \tag{47}$$

Similarly, we define the *Average Industry Cost of Outsourcing* as the mean friction faced by each industry across all of its inputs, or,

$$\tau_i^I = \frac{1}{N} \sum_k \tau_{ik} \tag{48}$$

Figure (15) shows the change in τ_k^O over time for each occupation k. The cost of outsourcing all but one task declined over this period, maybe by as much as 20%-30%. We see the

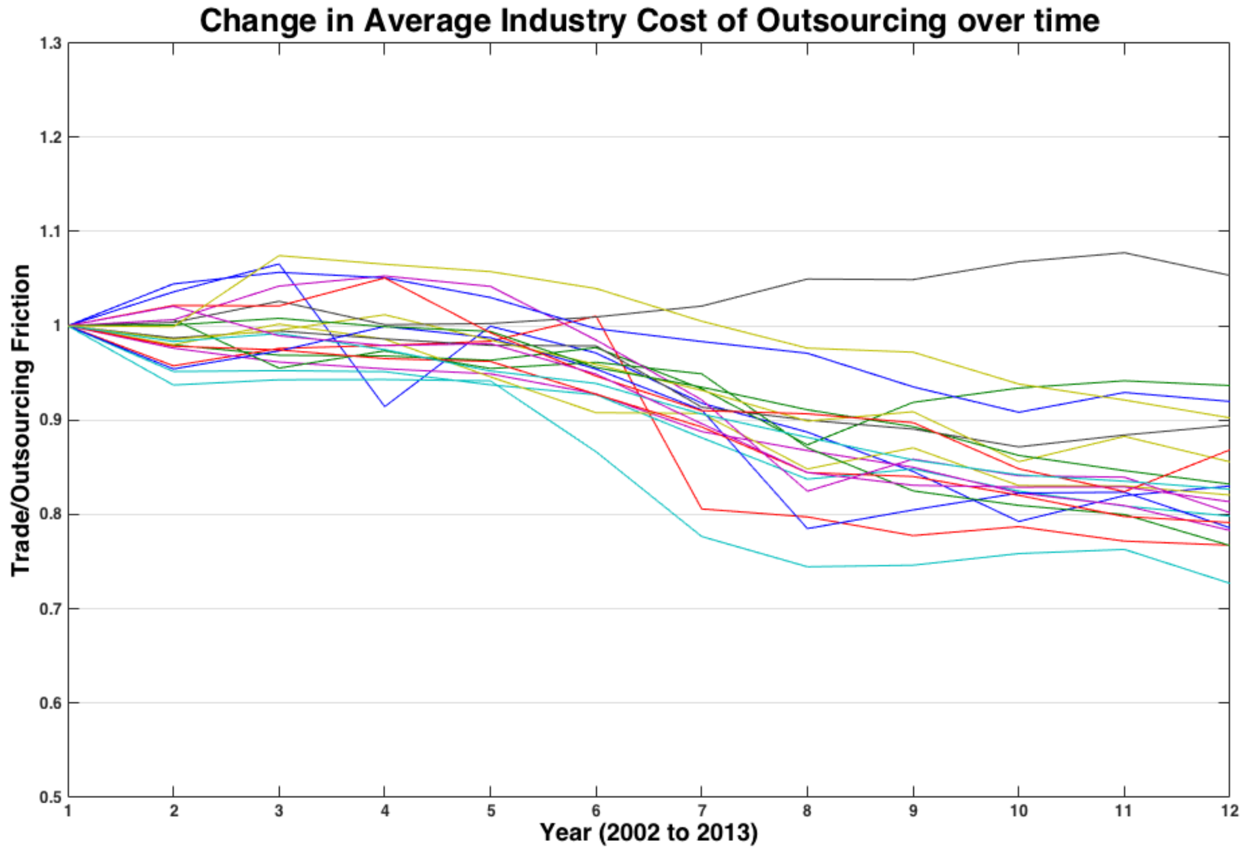

Figure 16: Change in the Average Industry Cost of Outsourcing τ_i^I over time

same trends when we look at τ_i^I. The average cost of outsourcing (in terms of strictly trade frictions) for almost every industry declined between 2002 and 2013. The key takeaway is that the incentives to purchase inputs from other industries, rather than produce them in-house with labour, have been increasing over time across the entire US economy. This decrease in frictions represents an increase in efficiency and output, but also has implications for labour demand and wage inequality. In particular, as workers are reshuffled across industries in response to changes in industry-occupation labour demand, we see shifts in the distribution wages which are due in part to sorting, as discussed in the empirical section, as well as the general equilibrium response of wages themselves to changes in demand and supply across industries with varying levels of productivity. The next section examines these dynamics by running a preliminary counter-factual exercise.

4.3 Counter-factual Exercises

In order to examine how changes in trade frictions have affected sorting patterns and wage inequality, we perform a simple counter-factual where we hold trade frictions for each industry pair at their 2002 level while allowing industry productivity and labour supply shocks to progress over time as measured by the calibrated model. In order to do this, we take the parameters and sequence of productivity and supply shocks from the calibrated model and recalculate the equilibrium wages, revenues, prices and labor demand in each period. This is a system of $2N + 2N^2$ equations with $2N + 2N^2$ unknowns. Our procedure makes an initial guess at the wages, then solves for labor supply, revenues and outsourcing share λ_{ik}, which gives us labor demand. We then define the function $F(w) = \sum_{i,k}(\ell^D_{ik}(w) - \ell^S_{ik}(w))^2$ where ℓ^D_{ik} and ℓ^S_{ik} represent labor demand and supply, respectively, for occupation k in industry i. The equilibrium system of wages w^* is such that the labor market clears for all occupations and industries, i.e.: $F(w^*) = 0$.

Change in...	**Baseline 2002-2013**	**Counterfactual 2002-2013**
Concentration	6.2%	-4.3%
Outsourcing Share	12.4%	-3.8%
Wage Variance	17.1%	11.6%

Given the counterfactual equilibrium in 2013, we can calculate the counterfactual changes in the objects of interest and compare them to the observed changes in the data. Table (**??**) shows the key objects of interest. The first row shows that the decrease in trade frictions between 2002 and 2013 drove all of the observed increase in occupational/industry concentration during this period. In fact, had trade frictions remained fixed over this time, increases in industry productivity would have actually decreased concentration (increased vertical integration). The second row similarly shows that the average outsourcing share, defined as the share of expenditure on inputs purchased from other industries (ie: the ratio of purchased materials to in-house labor), increased by 12.4% in the data. All of this increase was caused by the drop in trade frictions, as again without that change, the outsourcing share would have actually decreased by 3.8% over this period. Finally, the third row shows the change in wage variance (our measure of wage inequality). Roughly 1/3rd of the observed increase can be attributed to change in outsourcing costs. The remaining 2/3rds was due to changes in industry productivity and labour supply over this period. This is the key result from our exercise, and corroborates the empirical story above that much of the observed changes in wage inequality have been due not just to changes in wages, but actually to changes in outsourcing costs (trade frictions) which have led to a reshuffling of occupations towards their "home" industries. This occupation sorting, resulting from increased outsourcing, has been a significant driver of wage inequality not through changes in wages (which are driven primarily by productivity and labor supply) but through the mechanical reallocation of low-wage occupations to low-wage industries and vice versa via this outsourcing mechanism.

5 Conclusion

This paper proposes a simple model of firm specialization decisions which generates increased occupation sorting and wage dispersion in response to decreases in outsourcing costs. We show that this mechanism is consistent with observed trends in specialization, occupation concentration and wage dispersion in the data. Overall occupation employment concentration and sorting is growing over time, and variance decompositions of the trends in wage variance suggest that sorting accounts for a significant portion of changes in overall, between and within wage inequality. We examine this trend through the lense of the model, and estimate that about 1/3rd of the increase in inequality is due to decreases in intra-industry trade frictions, while the remaining 2/3rds are due to changes in technology and labor supply. In addition, the measured increases in occupational concentration are entirely due to decreased trade frictions, without which concentration would have actually decreased by 4.3%.

References

Autor, David, H, and David Dorn. 2013. "The growth of low-skill service jobs and the polarization of the US labor market." *The American Economic Review*, 103(5): 1553–1597.

Autor, David H, Lawrence F Katz, and Melissa S Kearney. 2008. "Trends in US wage inequality: Revising the revisionists." *The Review of economics and statistics*, 90(2): 300–323.

Baldwin, John R, Desmond Beckstead, and Richard E Caves. 2001. "Changes in the diversification of Canadian manufacturing firms (1973-1997): a move to specialization." *Statistics Canada, Analytical Studies-Micro-Economic Analysis Div. Series 11F0019MIE*, , (179).

Barth, Erling, Alex Bryson, James C Davis, and Richard Freeman. 2014. "It's where you work: increases in earnings dispersion across establishments and individuals in the US." National Bureau of Economic Research.

Caliendo, Lorenzo, and Esteban Rossi-Hansberg. 2012. "The Impact of Trade on Organization and Productivity." *The Quarterly Journal of Economics*, 1393–1467.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. 2015. "The Anatomy of French Production Hierarchies." *Journal of Political Economy*, 123(4): 809–852.

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

Chaney, Thomas. 2008. "Distorted gravity: the intensive and extensive margins of international trade." *The American Economic Review*, 98(4): 1707–1721.

Chaney, Thomas. 2014. "The Network Structure of International Trade." *American Economic Review*, 104(11): 3600–3634.

Chan, Mons. 2016. "Trade, Inequality, and the Nature of the Firm." Working Paper.

De Soyres, François. 2015. "Trade and Interdependence in International Networks." Working Paper.

Dube, Arindrajit, and Ethan Kaplan. 2010. "Does outsourcing reduce wages in the low-wage service occupations? Evidence from janitors and guards." *Industrial & labor relations review*, 63(2): 287–306.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, geography, and trade." *Econometrica*, 1741–1779.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2011. "An Anatomy of International Trade: Evidence from French Firms." *Econometrica*, 79(5): 1453–1498.

Faggio, Giulia, Kjell G Salvanes, and John Van Reenen. 2010. "The evolution of inequality in productivity and wages: panel data evidence." *Industrial and Corporate Change*, 19(6): 1919–1951.

Flood, Sarah, Miriam King, Steven J. Ruggles, and Robert Warren. 2016. "Integrated Public Use Microdata Series, Current Population Survey: Version 4.0 [dataset]." *Minneapolis: University of Minnesota.*

Grossman, Gene M, and Elhanan Helpman. 2015. "Growth, Trade, and Inequality." National Bureau of Economic Research.

Grossman, Gene M, and Esteban Rossi-Hansberg. 2008. "Trading Tasks: A Simple Theory of Offshoring." *American Economic Review*, 98(5): 593–629.

Grossman, Gene M, Elhanan Helpman, and Philipp Kircher. 2015. "Matching, Sorting, and the Distributional Effects of International Trade."

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii. 2012. "Identifying Equilibrium Models of Labor Market Sorting." National Bureau of Economic Research.

Handwerker, Elizabeth Weber, and James R Spletzer. 2013. "Occupational Concentration, Wages, and Growing Wage Inequality." Working Paper.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen J Redding. 2015. "Trade and inequality: From theory to estimation." National Bureau of Economic Research.

Holmes, Thomas J, and Julia Thornton Snider. 2011. "A Theory of Outsourcing and Wage Decline." *American Economic Journal: Microeconomics*, 38–59.

Itskhoki, Oleg, and Elhanan Helpman. 2015. "Trade Liberalization and Labor Market Dynamics with Heterogeneous Firms."

Kehoe, Timothy J, Kim J Ruhl, and Joseph B Steinberg. 2013. "Global imbalances and structural change in the United States." National Bureau of Economic Research.

Lemieux, Thomas. 2006. "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?" *The American Economic Review*, 461–498.

Yuskavage, Robert E, Erich H Strassner, and Gabriel W Medeiros. 2008. "Domestic Outsourcing and Imported Inputs in the US Economy: Insights from Integrated Economic Accounts."

Appendix A Mapping of Industry to Primary Occupation Group

In calibrating the equilibrium model of outsourcing and inter-industry trade, we need to identify a "primary" set of occupations which produce that industry's primary output separately from the set of occupations which are used to produce intermediate input tasks (i.e.: the primary output of other industries). We assume a one-to-one mapping consistent with the model and aggregate up to a set of 20 matched industry-occupation pairs, shown in table 3.

Table 3: Industry Occupation Mapping

Industry Title	**Occupation Title**
Farming, Forestry and Related	Farming, Fishing, and Forestry
Mining and Construction	Mining and Construction
Utilities and Telecom	Installation, Maintenance, and Repair
Manufacturing	Production
Wholesale, Retail,Trade and Real Estate	Sales and Related
Transportation, Warehousing and Waste Management	Transportation and Material Moving
Arts, Information and Media	Arts, Information and Media
Computer Systems, Design, Programming and Data Processing	Computer and,Mathematical Science
Finance and Insurance	Business and Financial Operations
Legal Services	Legal
Professional and Scientific Services	Architecture, Engineering and Science
Management of Companies and Enterprises	Management
Administrative and Support Services	Office Support and Cleaning Services
Educational Services	Education, Training, and Library
Health Care, Services and Hospitals	Healthcare, Practitioner and Technical
Nursing and Residential Care Facilities	Healthcare Support
Social Assistance	Community and Social Services
Amusements, Recreation, and Other Services	Personal Care and Service
Food Services and Drinking Places	Food Preparation and Serving Related
Government Services and Other	Protective Service & Others

Appendix B A Simple Model and Analysis

Appendix B shows the simpler version of the model, which has similar mechanisms with what we have in the full model in the main paper with the main difference being that here we set wages using union bargaining rather than household labour supply and labour market clearing. We have fully solved the simple model and analyze the dynamics of the model. The comparative statics analysis in the partial equilibrium is shown in Appendix C

B.1 Model

In order to better understand the dynamics surrounding our question, we build a model where heterogeneous industries (or representative firms) produce differentiated products using a set of intermediate inputs which can be purchased or made in-house using specialized occupations. We first describe the basic environment, do some simple comparative statics, and then show that the model is able to generate many of the stylized facts we see in the data regarding wage dispersion and occupational sorting.

B.2 Basic Environment

We consider an economy with a set I of industries, where each industry has a representative firm $i \in I$ which produces a unique output good or service, also indexed as $i \in I$, using firm i's primary labor and a set of intermediate inputs. These inputs can either be produced in-house by labor, or purchased from other firms. Firms sell their output to other firms as inputs, or on the final goods market at price P_i. Firms do not sell their own intermediate inputs. Workers are divided into occupations, which are indexed by the type of good which they produce. We suppose there is a single industry and single occupation per type of good. Notationally, L_{ij} is labor of type j used by firm type i, where $i, j \in I$.

B.3 The Firm

B.3.1 Production Technology

Firm i produces output using the following firm-specific production function:

$$Y_i = F_i(I_i, L_{ii}; A_i) \tag{49}$$

where I_i is a combination of intermediate input goods/services, L_{ii} is the quantity of primary occupation labor type i used by firm i, and A_i is a firm-specific general productivity parameter. Inputs I_i consists of a set of intermediate goods $\{\phi_{ij}\}_{j \neq i}$ using a CES aggregator

$$I_i = (\sum_{j \neq i} \alpha_{ij} \phi_{ij}^{\chi_i})^{\frac{1}{\chi_i}} \tag{50}$$

where χ_i determines the elasticity of substitution between inputs, and $\sum_{j \neq i} \alpha_{ij} = 1$. Note in particular that α_{ij} may be zero for some $j \neq i$ (ie: industries use some subset of the full set of goods as inputs into their own production).

Each ϕ_{ij} required by firm i can be produced in-house by hiring the appropriate type of labor L_{ij}, or purchased directly from firm j in quantity q_{ij}, or a combination, with $\phi_{i,j} = G(L_{ij}) + q_{ij}$.

B.3.2 Costs and Externalities

The cost of hiring labor type L_{ij} is

$$w_{ij} L_{ij} + M_i(L_{ij}) \tag{51}$$

where w_{ij} is the firm-occupation specific wage, and $M_i(L_{ij})$ is the firm-specific management cost of hiring intermediate labor L_{ij}. This could be thought of as any combination of agency or management costs due to monitoring, collective bargaining, enforcement etc which arise when managing intermediate labor. We allow some firms to be better at managing certain occupations than others.

The cost for firm i of purchasing q_{ij} of intermediate good j on the market is

$$P_j q_{ij} + c q_{ij} \tag{52}$$

where c is any sort of coordination, contracting, transportation or communication cost which affects the cost of inter-firm trade in inputs. We assume a common c for all firms, which in a dynamic model may be changing over time.

We also allow some firms to experience economies of scale from increased hiring of their primary labour type. This could be thought of as occupations or industries in which workers of a certain type have productive spillovers with other workers of their own type. For example, scientists may become more productive while working with other scientists. Software developers may build tools for their own use which also improve the productivity of their coworkers. In contrast, other occupations and industries may not benefit as much from these sorts of spillovers. When this externality is present, it works as if firms receive additional output as they increase employment above some minimum quantity required to operate: $e_i(L_{ii}, L_0)$. L_0 is lower bound for primary labour required to operate the industry or firm.

So for industries of this type, their primary labour type is crucial for production. and the more primary labour they hire, the greater the return.

B.3.3 The Firm's Problem

Given the above costs and production technologies, taking prices and wages as given, firm i solves the following profit maximization problem:

$$\max_{L_{ii},\{L_{ij},q_{ij}\}_{j\neq i}} P_i F_i\Big(\big(\sum_{j\neq i}\alpha_{ij}\phi_{ij}^{\xi_i}\big)^{\frac{1}{\xi_i}}, L_{ii}; A_i\Big) - w_{ii}L_{ii} - \sum_{j\neq i}\Big[w_{ij}L_{ij} + M_i(L_{ij}) + (P_j + c)q_{ij}\Big] \tag{53}$$

$$+ e_i(L_{ii}, L_0) \tag{54}$$

$$\begin{aligned} st. \quad & \phi_{ij} = G(L_{ij}) + q_{ij} \\ & L_{ii}, L_{ij}, q_{ij} \ \geq \ 0 \quad \forall\, j \neq i \end{aligned}$$

Firms solves their optimal labor and input demand as functions of wages and prices.

B.4 Wage Setting

Labour markets in our economy are not competitive. For each industry/occupation pair, there is a union or representative worker which sets wages in order to maximize their surplus, conditional on the industry's demand for labour and the worker's outside option. The union's problem is:

$$\max_{w_{ij}} (w_{ij} - b_{ij})L_{ij}(w_{ij}) \qquad \forall\, i, j \in I \tag{55}$$

where b_{ij} is the outside option (unemployment benefits) for occupation j workers in industry i. All workers' labor supply is perfectly elastic at the set wage (workers supply sufficient labor to meet industry labour demand at the union wage, and supply none at any other prices).

B.5 Market Equilibrium

In a market equilibrium, all firms are maximizing profits, workers are working at the optimal union wage, and goods markets clear. Here, we assume demand for a firm's output is the sum of demand from other firms plus some exogenous demand for consumption as a final good Q_{ij}. Goods market clearing requires that

$$Y_i = \sum_{j \neq i} q_{ji} + Q_i \qquad \forall\, i \in I \tag{56}$$

where q_{ji} is demand by firm j for good type i.

An equilibrium in this setting is a set of quantities $\{L_{ii}, L_{ij}, q_{ij}\}_{i,j \in I}$, and a set of prices $\{P_i, w_{ij}\}_{i,j \in I}$, such that given exogenous demand $\{Q_i\}_{i \in I}$, industries solve their maximization problems (53), unions maximize surplus (55), and markets clear (56).

B.6 Simple 2x2x2 Model

We solve and perform comparative statics with a basic version of the model, where we restrict the economy to two firms/industries, two occupations and two types of goods. As before, there is one firm per industry or output-type. Each industry produces its own output type using the other type of good as an input into its production process. The industry can either produce this input by hiring labor of the other type, or by purchasing the input directly from the other industry.

In order to make the model analytically tractable, we assume Cobb-Douglas production functions:

$$Y_i = A_i I_i^{\psi_i} L_{ii}^{\gamma_y}, \qquad \psi_i + \gamma_i \leq 1 \tag{57}$$

where I_i, the amount of input goods used by firm i is

$$I_i = G(L_{ij}) + q_{ij} = B_i L_{ij} + q_{ij}, \tag{58}$$

Here B_i represents the ease or productivity of hiring occupation j to produce in-house relative to purchasing the input on the open market. $B_i > 1$ implies that the industry may require a more specialized version of the input than if $B_i \leq 1$, which would imply that non-custom products bought on the market are just as productive.

We restrict industry i to produce output using a fixed labor-intermediate ratio $L_{ii}/I_i = \theta_i$. In this paper, we fix θ_i exogenously which can be interpreted as a short term technological constraint. The model can be easily extended such that industries/firms choose their optimal input ratio given union wage setting behaviour, which can be interpreted as industries adjusting production technology in the long run. However, in the short run, firm i has to use a $(I_i = 1/\theta_i, L_{ii} = 1)$ pair to produce $Y_i = \frac{A_i}{\theta^{\psi_i}}$ unit of output. This restriction is functionally equivalent to using a Leonteif-style production technology.

We specify management costs as

$$M_i(L_{ij}) = L_{ij}^{\alpha_i}, \qquad \alpha_i > 1 \tag{59}$$

This specification has several implications. First, the convex management cost and linear input purchase cost implies that neither firm will entirely divest itself of input labor. There will always be some optimal level of intermediate labor, above which all additional input is purchased. This is not an entirely unreasonable assumption. Consider that many businesses may still employ a lawyer in-house despite contracting most of their required law service inputs to external firms. In-house labor with both task and firm-specific human capital, can coordinate and direct the contracting of outsourced labor. However, extensions of the model (currently a work in progress) allow for corner solutions on both dimensions. We expect the results will be qualitatively similar. Additionally, if the industry's primary labour exhibits positive externalities, then

$$e_i(L_{ii}, L_0) = \xi_i \log(L_{ii} - L_0) \tag{60}$$

We will refer to industry-occupation pairs with positive externalities as type-1 industries, and the ones without as type-2 industries. We will analyze the problems of both types of industry, then introduce market equilibria in which there is between and within-type trading.

B.6.1 Type 1 Industry's Problem

An industry i of type 1 has the following problem:

$$\begin{aligned}\max_{L_{ii},L_{ij},q_{ij}} \quad & P_i A_i I_i^{\psi_i} L_{11}^{\gamma_i} - w_{ii}L_{ii} - \left[w_{ij}L_{ij} + L_{ij}^{\alpha_i} + (P_j + c)q_{ij}\right] + \xi_i \log(L_{ii} - L_0) \\ st. \quad & I_i = B_i L_{ij} + q_{ij} \\ & L_{ii}/I_i = \theta_i \\ & L_{ii}, L_{ij}, q_{ij} \ \geq \ 0\end{aligned}$$

To simplify the algebra, we assume a CRS production technology: $\psi_i + \gamma_i = 1$. When $\psi_i + \gamma_i < 1$, the result is qualitatively the same. Solving the industry's problem results in the following labour demand as functions of prices, wages and parameters.

$$L_{ii}^D = L_0 + \frac{\xi_i}{w_{ii} - (\frac{P_i A_i}{\theta_i^{\psi_i}} - \frac{c+P_j}{\theta_i})} \tag{61}$$

$$L_{ij}^D = \left(\frac{B_i(c+P_j) - w_{ij}}{\alpha_i}\right)^{\frac{1}{\alpha_i - 1}} \tag{62}$$

Note that when wages are held constant, L_{ii} is increasing in the outsourcing cost c, while L_{ij} is decreasing in c. This is intuitive: When the outsourcing cost declines, industries purchase more intermediate input and produce less in-house. Also because of the short term technological restriction on the primary labour-intermediate ratio, industries additionally hire more primary labor. So when outsourcing costs decrease, the labour share for primary industry labor share increases, increasing occupational concentration/specialization. This is what we refer as the mechanical sorting and labor demand effects.

Before we show the wage effect of the changes in outsourcing costs, it is useful to see what are the conditions for industries to have an optimal solution to their profit maximization problem (ie: a unique labour and input quantity demand schedule as a function of prices and parameters). The optimality condition can be clearly interpreted as shown in 17. The red line in the graph shows the marginal revenue when a firm uses a labor input pair ($L_{ii} = 1, I_i = \frac{1}{\theta_i}$) to produce $\frac{A_i}{\theta_i^{\psi_i}}$ units of output. So the firms marginal revenue from the production using this labor input pair is

$$MR = \frac{P_i A_i}{\theta_i^{\psi_i}} + \frac{\xi_i}{L_{ii} - L_0}$$

where L_{ii} is how much primary labor that the firm has employed so far. Notice that the MR

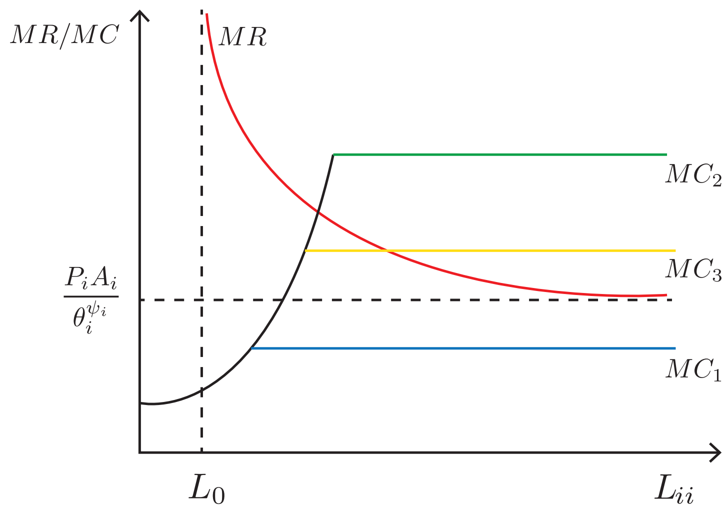

Figure 17

is decreasing in L_{ii} and converges to $\frac{P_iA_i}{\theta_i^{\psi_i}}$ as L_{ii} increases.

The marginal cost for the production using this labor input pair is slightly more complicated. The management cost of the secondary labor is convex in the amount of labor (L_{ij}), and the secondary labor and the purchased inputs (q_{ij}) are perfect substitutes in production. This means the industry will always produce the intermediate inputs in-house initially by hiring secondary labor until the marginal cost of hiring L_{ij} is equal to the outsourcing cost ($P_j + c$). Then the firm will switch to buying the inputs. The marginal cost is thus

$$MC = \begin{cases} w_{ii} + \frac{w_{ij}}{\theta_i B_i} + \frac{\alpha_i (L_{ij})^{\alpha_i - 1}}{\theta_i B_i} & \text{if } L_{ij} \leq \left(\frac{B_i(c+P_j) - w_{ij}}{\alpha_i}\right)^{\frac{1}{\alpha_i - 1}} = L_{ij}^D \\ w_{ii} + \frac{P_j + c}{\theta_i} & \text{otherwise} \end{cases}$$

Where L_{ij} is the unit of secondary labor employed so far.

Figure 17 shows three possible conditions. Firstly, when $\frac{P_iA_i}{\theta_i^{\psi_i}} > w_{ii} + \frac{P_j+c}{\theta_i}$ (blue line), industries do not have an optimal level of demand for labour, since $MR > MC$ for all production levels, leading to infinite output. Secondly, when $\frac{P_iA_i}{\theta_i^{\psi_i}} > w_{ii} + \frac{P_j+c}{\theta_i}$, and $w_{ii} +$

$\frac{P_j+c}{\theta_i} \geq \frac{P_i A_i}{\theta_i^{\psi_i}} + \frac{\xi_i}{\frac{1}{\theta_i B_i} L_{ij}^D - L_0}$ (green line), the industry's optimal strategy is to not outsource and produce everything in-house. Thirdly, when $\frac{P_i A_i}{\theta_i^{\psi_i}} > w_{ii} + \frac{P_j+c}{\theta_i}$, and $w_{ii} + \frac{P_j+c}{\theta_i} < \frac{P_i A_i}{\theta_i^{\psi_i}} + \frac{\xi_i}{\frac{1}{\theta_i B_i} L_{ij}^D - L_0}$ (yellow line), the industry's optimal strategy is to employ L_{ij}^D units of secondary labor, and outsource the rest: $q_{ij}^D = \frac{L_{ii}^D}{\theta_i} - B_i L_{ij}^D$

Given industry optimal labour demand as a function of price and wages, in equilibrium, unions set wages such that surplus going to labour is maximized. The unions' problems are:

$$\max_{w_{ii}} \quad (w_{ii} - b_{ii}) L_{ii}^D(w_{ii}) \tag{63}$$

$$\max_{w_{ij}} \quad (w_{ij} - b_{ij}) L_{ij}^D(w_{ij}) \tag{64}$$

Which results in the unions' optimal wage setting:

$$w_{ii} = \left[\Big(\frac{P_i A_i}{\theta_i^{\psi_i}} - \frac{c+P_j}{\theta_i} - b_{ii} \Big) \xi_i / L_0 \right]^{\frac{1}{2}} + \frac{P_i A_i}{\theta_i^{\psi_i}} - \frac{c+P_j}{\theta_i} \tag{65}$$

$$w_{ij} = \frac{(\alpha_i - 1)(c+P_j) B_i + b_{ij}}{\alpha_i} \tag{66}$$

It is clear from the result above that when the outsourcing cost c decreases, the primary labour's wage w_{ii} increases, and the secondary labour wage w_{ij} decreases. If the primary occupation in the industry earns higher wages than the secondary occupation, for example, the law services industry which hires lawyers and janitors, then decreasing outsourcing costs can lead to wage polarization within the industry. On the other hand, if the primary occupation's wage is relatively lower than the secondary occupation's wage (as may be true in the janitorial services industry), then the decreasing outsourcing cost can decrease within-industry wage dispersion.

What about the industry's demand response to this wage change? Plugging (65) and

(66) back into the industry labour and input demand functions, we get:

$$L_{ii}^{D} = L_0 + \left(\frac{\xi_i L_0}{\frac{P_i A_i}{\theta_i^{\psi_i}} - \frac{c+P_j}{\theta_i} - b_{ii}} \right)^{\frac{1}{2}} \tag{67}$$

$$L_{ij}^{D} = \left(\frac{B_i(c+P_j) - b_{ij}}{\alpha_i^2} \right)^{\frac{1}{\alpha_i - 1}} \tag{68}$$

$$q_{ij} = \max\{ \frac{L_{ii}}{\theta_i} - B_i L_{ij}, 0 \} \tag{69}$$

So the total effect (wage effect + demand effect + mechanical sorting) leads to a decrease in both primary and secondary labor employment in response to the decrease in the outsourcing cost, with the magnitudes depending on the parameters. So far we have examined the individual industry problem, taking prices P_i and P_j as given. We will endogenize prices in a market equilibrium in a later section. Now we can introduce the conditions for optimality more formally:

Suppose that the following two assumptions are satisfied:

1. $\dfrac{P_i A_i}{\theta_i^{\psi_i}} - \dfrac{c + P_j}{\theta_i} > b_{ii}$
2. $\dfrac{P_i A_i}{\theta_i^{\psi_i}} - \dfrac{c + P_j}{\theta_i} < b_{ii} + \dfrac{\xi_i L_0}{\theta_i^2 B_i^2 \left(\frac{B_i(c+P_j) - b_{ij}}{\alpha_i^2} \right)^{\frac{2}{\alpha_i - 1}}}$

Then there exists a unique solution for firm i's problem with non-negative profits where the optimal outsourcing quantity is greater than zero.

Proof. Given the monotonicity of the marginal revenue and the marginal cost function, the theorem follows trivially given figure (17), and equations (65) and (68). □

B.6.2 Type 2 Industry's Problem

Now let's assume firm i is a representative firm in a type 2 industry. A Type 2 firm's problem is very similar with a type 1 firm, except that there is no externality effect from the primary occupation.

Firm i solves the following problem:

$$\max_{L_{ii},L_{ij},q_{ij}} P_i A_i I_i^{\psi_i} L_{11}^{\gamma_i} - w_{ii}L_{ii} - \left[w_{ij}L_{ij} + L_{ij}^{\alpha_i} + (P_j + c)q_{ij}\right]$$
$$st. \quad I_i = B_i L_{ij} + q_{ij}$$
$$L_{ii}/I_i = \theta_i$$
$$L_{ii}, L_{ij}, q_{ij} \ \geq \ 0$$

Note that here we made a slight modification: $\psi_i + \gamma_i < 1$. This is because that with CRS technology, there would be no optimal labour and input quantity demanded for firm i, similar to the arguments demonstrated in figure 17. Because of the similarity with the previous section, here we omit most of the analysis and only show the results.

The type 2 firm's labor demand is:

$$L_{ii}^D = \left(\frac{\left(\frac{c+P_j}{\theta_i} + w_{ii}\right)\theta_i^{\psi_i}}{P_i A_i(\psi_i + \gamma_i)}\right)^{\frac{1}{\psi_i+\gamma_i-1}} \tag{70}$$
$$L_{ij}^D = \left(\frac{B_i(c + P_j) - w_{ij}}{\alpha_i}\right)^{\frac{1}{\alpha_i-1}} \tag{71}$$

Similar with the type one firm's problem, when wages and prices are fixed, decreasing outsourcing cost results in an increase in the primary labor demand, and decrease in the secondary labor demand. The reason is same as the previous case. Mechanical sorting and the demand effect increases the primary labor share when outsourcing cost decreases.

The unions optimal wage settings are:

$$w_{ii} = \frac{(1 - \psi_i - \gamma_i)\frac{c+P_j}{\theta_i} + b_{ii}}{\psi_i + \gamma_i} \tag{72}$$
$$w_{ij} = \frac{(\alpha_i - 1)(c + P_j)B_i + b_{ij}}{\alpha_i} \tag{73}$$

The union wage setting results are different from the type 1 firm's case. When outsourcing costs decrease, both the primary and secondary labor wages decrease. Whether wage dispersion rises or shrinks depends on the parameter choices and initial relative employment share. The reason that the primary labor's wage decreases here (increases in the type 1 firm's case) is as follows: because of the DRS production technology and no positive externality from the primary labor, firm's demand elasticity for primary labor is increasing as c decreases. When the union maximizes the labour surplus, decreasing the wage will result in a relatively large labour demand response so that total labour surplus increases. This is the opposite result from the type 1 case, where industry primary labour demand becomes less price elastic as c decreases.

Given the union's wage setting, firms labor and input demands are:

$$L_{ii}^D = \left(\frac{\left(\frac{c+P_j}{\theta_i} + b_{ii}\right)\theta_i^{\psi_i}}{P_i A_i (\psi_i + \gamma_i)^2} \right)^{\frac{1}{\psi_i+\gamma_i-1}} \tag{74}$$

$$L_{ij}^D = \left(\frac{B_i(c+P_j) - b_{ij}}{\alpha_i^2} \right)^{\frac{1}{\alpha_i-1}} \tag{75}$$

$$q_{ij} = \max\{\frac{L_{ii}}{\theta_i} - B_i L_{ij}, 0\} \tag{76}$$

So the total effect (wage effect + demand effect + mechanical sorting) leads to an increase in primary labor and decrease in secondary labor in response to the decrease in the outsourcing cost.Similar with the type 1 firm's problem, now we can introduce the conditions for optimality:

Suppose that the following assumption is satisfied:

$$\frac{P_j + c + b_{ii}\theta_i}{P_i A_i (\psi_i + \gamma_i)^2 \theta^{1-\psi_i}} < (\theta_i B_i)^{\psi_i+\gamma_i-1} \left(\frac{B_i(c+P_j) - b_{ij}}{\alpha_i^2} \right)^{\frac{\psi_i+\gamma_i-1}{\alpha_i-1}}$$

Then there exist an unique optimum for firm i where the optimal outsourcing quantity is greater than zero.

Proof. Given the monotonicity of the marginal revenue and the marginal cost function, the theorem follows trivially given (72) and (76). □

Appendix C Comparative Statics in Partial Equilibrium

Our main mechanism of interest is how wages and labor allocations within and between firms adjust in response to changes in the cost of outsourcing, c. Our story is that decreases in this outsourcing cost increases firm specialization and occupation sorting, ie: that firms respond on the intensive margin to decreases in outsourcing costs by decreasing the quantity of employed intermediate occupation labor, and (in a model where we allow corner solutions for labor demand), respond on the extensive margin by decreasing the cardinality of the set of intermediate occupation types they employ. We show that our model, which is at its heart a very general extension of the basic model of firm behavior used in previous literature, generates exactly these results, even in a closed economy framework with fixed prices. We also show that decreasing outsourcing frictions generates significant wage changes in our model, which also contribute to changes in the wage distribution.

To do this, we solve a partial equilibrium version of our model with two industries and two occupations as described above. Here we assume both industries are price takers with respect to their output goods (we discuss endogenous output prices in the next section). We allow firms to trade with each other (restricting output to be at least as great as the demand for their product) as well as sell any surplus output to an outside market at the fixed prices.

We consider three different economies in order to explore how changes in outsourcing frictions affect wages, sorting and labour in our model. In order to graph these responses, we choose a set of reasonable parameters for the model and use our closed form solutions to show how the equilibrium prices and quantities respond to changes in c - the outsourcing friction. In all of the following exercises, the industries are symmetric with the exception of the industry-level technology parameters A_i and B_i. Specifically, we choose $A_1 > A_2$, implying that industry 1 is generally more productive than industry 2. We also set $B_1 < 1 < B_2$, so that hiring occupation 1 to produce input services is more productive than outsourcing from the other firm (without considering cost), while hiring occupation 2 is less productive. One way to think of this is that it represents differences in how generalizable an input is. Some inputs are very specific to the industry or firm (such as custom parts or software), while others are very general (such as janitorial services). So our setting is similar to our original example, with one high tech software firm and one janitorial firm.

C.1 Occupation Sorting and Wages in Partial Equilibrium

Figure 18 shows an economy with two CRS-technology firms with primary occupation externalities, where wages are set by workers/unions as described above. This is our baseline model, since it is able to generate most of the dynamics we see in the aggregate wage data at the industry and occupation level.

All panels share the same x-axis, which represents changes in the outsourcing cost c. The top left panel shows changes in total wage variance as well as its between industry and within industry, between occupation components (which matches what we measure in the OES data in the empirical section)[25]. The top right panel shows changes in the Industry Herfindahl Index[26] which is a measure of the concentration of occupations employed within an industry or firm. We calculate the Herfindahl for each industry, and the weighted mean Herfindahl for the entire economy.

The next two panels show changes in labour demand and equilibrium wages as c changes. The bottom left panel ("Labour Ratios") shows the ratio of primary industry employment to input labour employment for both occupations (ie: L_{ii}/Lij) as well as the ratio of total employment in Firm 1 to Firm 2. If these ratios increase, it means that the occupation is increasingly sorted into its primary industry. The bottom right panel shows average labour wages for both occupations and both firms as well as the overall mean wage.

As outsourcing costs decrease (moving left on the x-axis), we get a number of effects. First, the wages for primary occupation labour in both industries increases, while wages for input labour of both types decrease. Demand for input labour decreases in both industries, and the ratio of primary to input labour grows as both industries outsource input labour. Demand for primary labour remains relatively constant. This causes the Herfindahl Index to increase in both industries as they become increasingly specialized in their primary output/occupation. One interesting result is that more productive firms have a higher Herfindahl, and so are more specialized than firms with lower general productivity.

The graph showing changes in wage variance deserves special attention, since that is the primary object of interest in our overall analysis of occupation sorting and wage inequality. Our model captures all three of the mechanisms we describe earlier in the paper. First, decreases in outsourcing costs lead to changes in the relative employment of primary and input labour in all industries, leading to mechanical changes in the within and between components of wage variance. Second, as purchasing inputs becomes relatively cheaper,

[25] Total wage variance is decomposed as follows:

$$\text{Var of log wages} = \mathbb{E}_{io}[(w_{io} - \bar{w})^2] = \mathbb{E}_i[(\bar{w}_i - \bar{w})^2] + \mathbb{E}_{io}[(w_{io} - \bar{w}_i)^2]$$

where $\bar{w}$ is the overall mean wage, $\bar{w}_i$ is the mean wage in industry i, and w_{io} is the wage for occupation o in industry i.

[26] We define the Industry Herfindahl Index for industry i as $H_i = \sum_{o \in O_i} s_{io}^2$, where s_{io} is the employment share of occupation o in industry i. A Herfindahl Index of 1 implies that an industry employs only a single occupation, while a Herfindahl close to 0 implies significant heterogeneity in within-industry occupation employment.

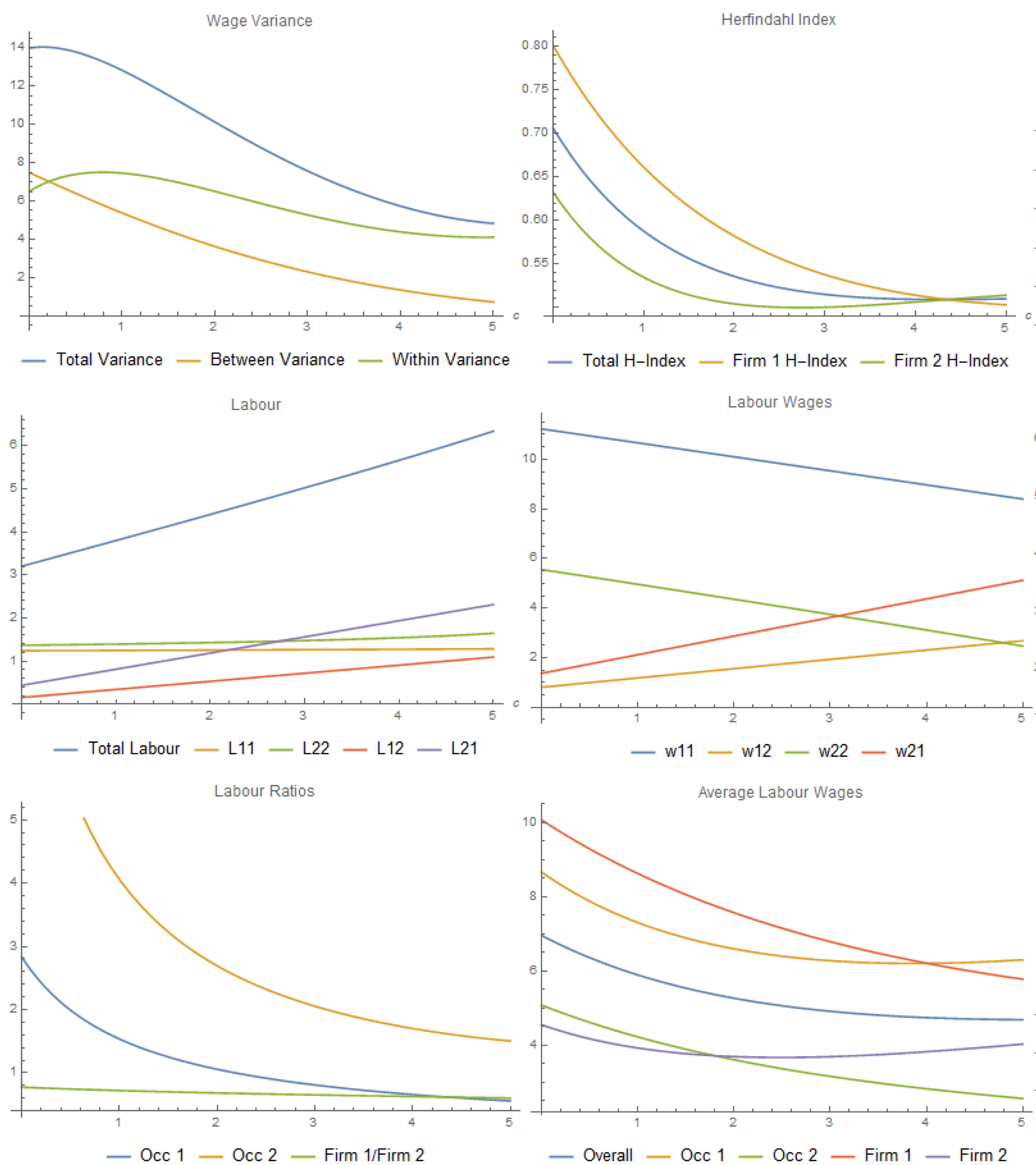

Figure 18: Economy with two type-1 industries and wage setting.

overall demand for labour decreases at different rates across occupations, leading to changes in overall wage variance. Thirdly, changes in outsourcing costs lead to changes in wages for both occupations. The overall effect is an increase in total variance, coming mostly from growth in between industry variance, while within industry variance remains fairly constant. This is similar to what we see in the data (see figure 1).

Much like the counterfactual variance decomposition we perform using the OES data in figures 1 and 4, we can determine whether this change in wage inequality is being driven primarily by changes in wages, or purely changes in the composition of occupations within and between industries. Unlike the empirical exercise, however, doing so in our model allows us to account for equilibrium effects on wages.

Figure 19 shows exactly the same economy, with the exception that we hold wages fixed when solving for labour and input demand. There are several main differences in this economy. The main difference is that since wages are held fixed, the only changes in wage inequality come from occupation sorting and changes in aggregate labour demand. This sorting effect drives between industry wage inequality up (due to increasing differences in average industry wage - see the last panel), while driving within firm inequality down. The overall effect is that total wage variance changes only slightly (due entirely to the change in aggregate labour demand), since the between and within effects from sorting mechanically cancel each other out. Thus, in our baseline model, it's clear that the increase in wage variance from a decrease in outsourcing frictions is due primarily to wages increasing within industry wage variance, rather than between industry variance.

One limitation of our baseline setting, where both industries have CRS technology with the primary occupation externality (type-1 industries), is that wages for primary industry occupations always increase as c decreases. However in the data, we show that this is true for some occupations (typically high skill occupations) but not all. In order to allow for both possibilities in our model, we solve an equilibrium where one industry is as before, while the other is a decreasing returns to scale technology industry with no externality effect from the primary occupation (a type-2 industry)[27]. This gives us a setting in which one firm has increasing labour demand elasticity as $c \downarrow$ (leading to decreasing wages) and the other has decreasing demand elasticity (increasing wages). Figure 20 shows similar patterns in wage variance as before (increasing overall variance driven primarily by increased between variance). The Herfindahl for firm 2 is decreasing since it initially employs more input labour

[27]We also solve an equilibrium with two type-2 industries. However, it provides less useful results, since it is unable to generate the kinds of wage and labour dynamics we see in the data.

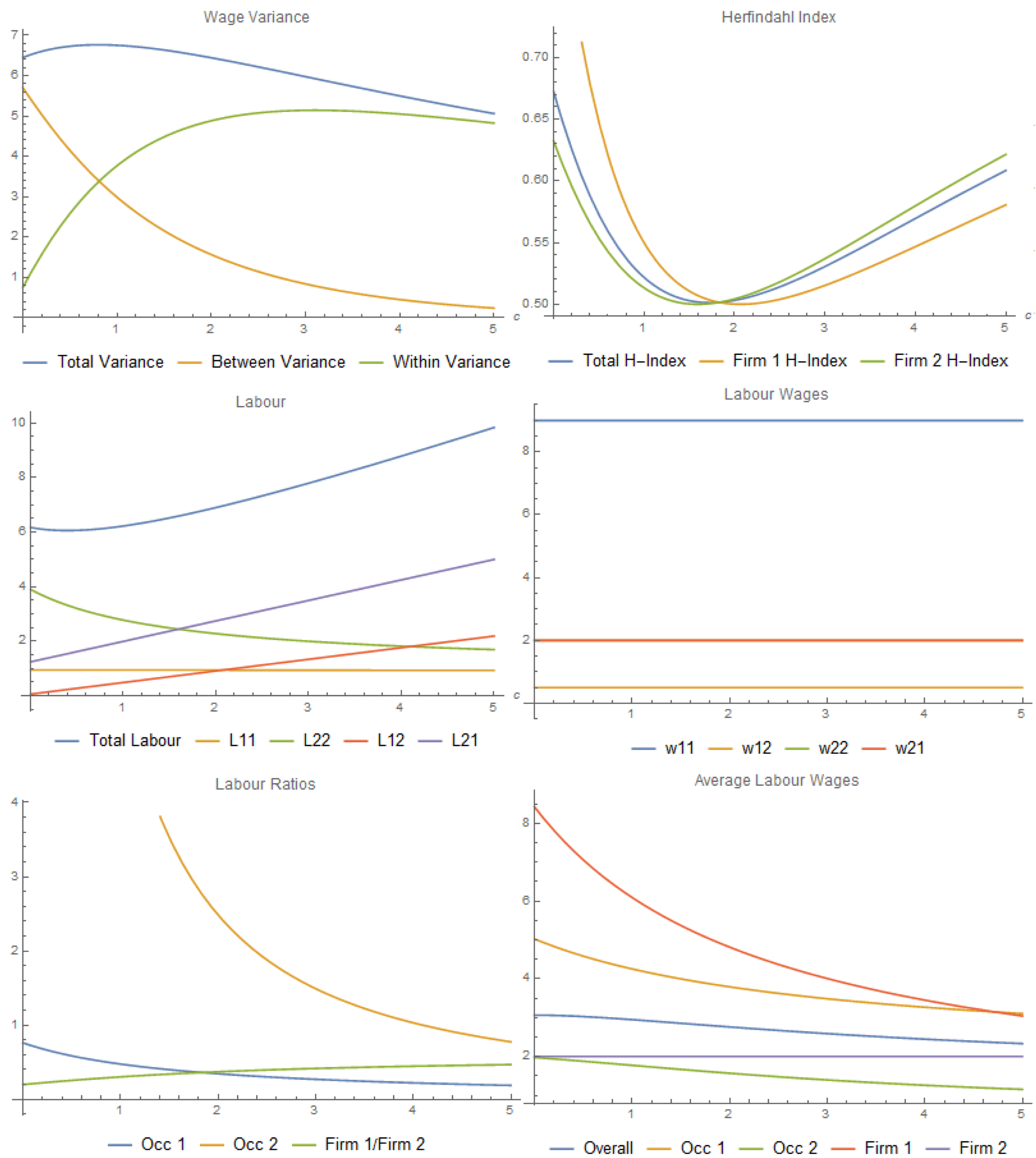

Figure 19: Economy with two type-1 industries and fixed wages.

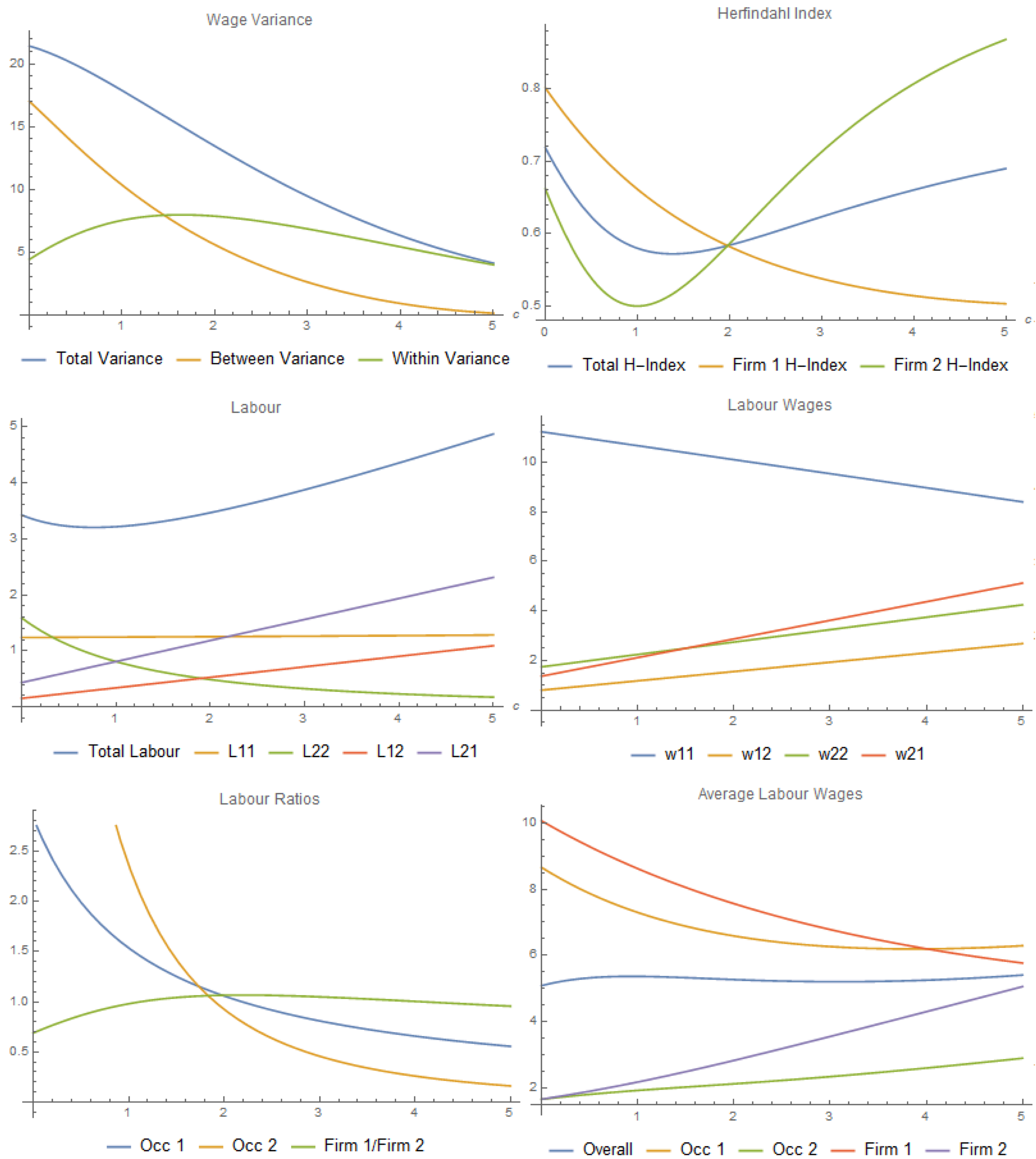

Figure 20: Economy with one type-1 industry, one type-2 industry and wage setting.

than its primary occupation. The Herfindahl begins to increase again once this ratio flips due to outsourcing. This is not unexpected, and also matches what we see in the data, as not every industry is increasing in the Herfindahl index, even though they may be outsourcing and sorting as in our model economy.

So, our very simple model is able to generate the relationships and responses to changes in outsourcing costs which are consistent with the basic mechanism we established in the introduction, even in a static, symmetric partial equilibrium setting with two industries/firms.

C.2 Market Equilibrium with Endogenous Prices

We can extend the analysis of our model further by allowing for prices to be set in equilibrium. We do so by fixing an exogenous level of demand for each industry's output, representing market demand for output as final consumption goods. Total demand for an industry's output is then this exogenous market demand plus the demand for that output as an input into the other industry's production process. We assume that industries can not sell their intermediate inputs. In a $2 \times 2 \times 2$ model, the following two market clearing conditions are satisfied:

$$A_i(B_iL_{ij} + q_{ij})^{\psi_i} L_{ii}^{\gamma_i} = \bar{Q}_i + q_{ji} \tag{77}$$

$$A_j(B_jL_{ji} + q_{ji})^{\psi_j} L_{jj}^{\gamma_j} = \bar{Q}_j + q_{ij} \tag{78}$$

$\bar{Q}_i$ and $\bar{Q}_j$ are exogenous demand for goods i and j. In the previous analysis, we calculated all the optimal quantities as functions of prices and parameters: $L_{ii}(P_i, P_j; \Theta)$, $L_{ij}(P_i, P_j; \Theta)$, $q_{ij}(P_i, P_j; \Theta)$, where Θ is the set of model parameters. Plugging these optimal quantities into the market clearing conditions (77) and (78) allows us to pin down all the prices in terms of the model parameters: $P_i = P_i(\Theta)$, $P_j = P_j(\Theta)$. Given these equilibrium prices, all the equilibrium quantities $(L_{ii}, L_{ij}, L_{ji}, L_{jj}, q_{ij}, q_{ji})$ and equilibrium wages $(w_{ii}, w_{ij}, w_{ji}, w_{jj})$ can also be solved purely as functions of model parameters.

Given the current assumptions on parameters and functional forms, getting closed form solutions for P_i and P_j may not be possible. However, even without closed form solutions, we can still do some qualitative analysis using numerical methods. We are currently working on fully simulating a generalized version of the model. Our initial results are very similar to the equilibrium with fixed prices above, though how sensitive the model is to price effects depends on the choice of parameters.

C.3 Empirical Implications

The analysis of the model provides several predictions which we can test in the data. Specifically, decreases in common outsourcing costs due to technological or policy change should be associated with increasing Employment Herfindahl indexes within firms, industries and across the economy as a whole. Additionally, we should see increasing between-industry wage variance and declining within wage variance due to compositional change/sorting, and increases in within industry wage variance due to changes in wages. We could also potentially test the prediction that more productive firms are less specialized, or that changes in outsourcing are associated with overall decreases in labor demand or increases in aggregate output. Indeed, these are exactly the empirical facts we see in the data.

Appendix D Proofs of Results in Comparative Statics Section