New Stylized Facts on Occupational Employment and Their Implications: Evidence from Consistent Employment Data*

Myungkyu Shim† Hee-Seung Yang‡

August 12, 2016

Abstract

The business cycle properties of occupational employment have not yet been extensively explored because of inconsistencies in the aggregate employment series by occupation. Using consistent aggregate hours data constructed through the method of "conversion factors," which was developed by the U.S. Census Bureau, we provide new empirical facts on the cyclical behaviors of occupational employment and discuss their implications. First, employment of the middle-skill occupation group is negatively affected by a technology shock, while those of high-skill and low-skill groups are positively correlated with it. Second, it is the middle-skill group that experiences the largest decline in employment volatility after the mid-1980s. Last, recessions since the 1980s have heterogenous impacts on different occupations, defining the characteristics of each recession. We further discuss the value of having consistent employment data in studies of business cycles.

JEL classification: C82, E24, E32

Keywords: Business cycle property; Occupational employment; Consistent data; Conversion factor; VAR; Employment volatility

*First Draft: May 2012. A previous version of the paper was circulated under the titles "Business Cycle Properties of Job Polarization Using Consistent Occupational Data" or "Unveiling Business Cycle Properties of Occupational Employment using Consistent Data of Occupations." The authors gratefully acknowledge the constructive comments made by three anonymous referees and the editor. The authors thank David Autor, Tomaz Cajner, David Dorn, Youjin Hahn, Pushkar Maitra, Bruce Preston, Valerie Ramey, and Russell Smyth for their helpful comments. Yu Jung Whang provided excellent research assistance. Hee-Seung Yang gratefully acknowledges the financial support from Monash University.

†Shanghai University of Finance and Economics and Sogang University. Email: audrb917@gmail.com.

‡Monash University. Email: heeseung.yang@monash.edu.

1 Introduction

One of the fascinating topics in the labor economics and macroeconomics fields is the causes and consequences of structural changes in the labor market. In particular, "job polarization," that is, the increase in job opportunities at the two ends of the skill distribution and the decline in opportunities in the middle, attracts researchers to study the properties of occupational employment (see Acemoglu and Autor (2011); Autor, Katz, and Krueger (1998); Cortes (2016); Autor and Dorn (2013); and Jung and Mercenier (2014)). Interestingly, while the long-run properties of occupational employment (i.e., employment of high-skill, middle-skill, and low-skill occupation groups) have been well-documented in the literature, the business cycle properties of occupational employment have not yet been extensively explored.[1] Jaimovich and Siu (2014), Foote and Ryan (2014), Gaggl and Kaufmann (2015), and Smith (2013) are few exceptions that we are aware of. One major obstacle in the study of short-run fluctuations is the presence of inconsistencies in aggregate employment series by occupation because occupation classifications have changed over time and the changes in occupation codes overlap key moments of the business cycle such as recessions (see Dorn (2009) and Foote and Ryan (2014)). Despite such an obstacle, studying the business cycle properties of occupational employment is important because employment fluctuations affect the aggregate economy in many dimensions. For instance, they can affect the welfare cost of business cycles; when the asset market is incomplete, the welfare cost of business cycles depends heavily on labor income dynamics. As wage rates and skill premia are almost acyclical, employment fluctuations have sizable impacts on the welfare costs across occupation groups (Castro and Coen-Pirani (2008); and Shim and Yang (2015)).

Our paper attempts to fill this gap in the literature in two ways. First, we provide consistent aggregate employment data for different occupation groups by using the method of "conversion factors." While we are not the first to use this method (for example, Beaudry, Green, and Sand (2015) and Lefter and Sand (2011) employed the same method, but they analyzed long-run trends only), we are the first to apply the method to study short-run behaviors of occupational employment. Second, using the consistent aggregate hours data by occupation, we provide several empirical regularities of occupational

[1] High-skill (non-routine cognitive) occupations include "Managers," "Professionals," and "Technicians." Middle-skill (routine) occupations include "Sales," "Office and administration," "Production, crafts, and repair," and "Operators, fabricators, and laborers." Low-skill (non-routine manual) occupations include "Protective services," "Food prep, building and grounds cleaning," and "Personal care and personal services." See Acemoglu and Autor (2011) or Autor (2010) for details.

employment at the business cycle frequency that have not been documented before. This is particularly important when we need to identify the heterogenous impacts of business cycles on different occupation groups. For instance, for implementing a labor market policy that attempts to help the workers who are mostly affected by a recession, we need information on the dynamics of the labor market, that is not contaminated by inconsistencies in data.

We first address the inconsistency issue by comparing aggregate employment and total hours series for each occupation group by using the following two methods: (1) occ1990dd classification and (2) conversion factors. Details of these methodologies are discussed in Section 2. By comparing aggregate hours variables obtained through each method, as presented in Section 3, we find that the method of conversion factors provides aggregate hours data without any discontinuities during the sample period. In contrast, the occ1990dd data and the raw data, which are constructed without any methods, exhibit breaks in aggregate data when occupation codes change. In principle, the occ1990dd classification is neither intended to construct aggregate employment series nor to be used for business cycle analysis; however, the method has been employed by several papers such as Foote and Ryan (2014) and Gaggl and Kaufmann (2015) that study business cycle properties of occupational employment. They circumvent the inconsistency issue by using growth rates of employment variables instead of using levels. In this sense, the data constructed through conversion factors enable researchers to conveniently use levels of aggregate employment series without discontinuities so that it can enrich our understanding of employment fluctuations.

We then present three stylized facts about cyclical properties of occupational employment, which have never been documented before: (1) the heterogenous responses of occupational employment to technology shocks, (2) the degree of changes in volatility of hours worked of each occupation group since the mid-1980s, and (3) the asymmetric effects of recessionary episodes on occupational employment. We first show, using a vector autoregression (VAR, henceforth) with a long-run restriction à-la Galí (1999), that it is the middle-skill occupation group that is negatively affected by the technology shock; the high-skill and low-skill groups are instead positively related to the technology shock. This result suggests that the finding by Galí (1999) that hours negatively react to the identified technology shock might be driven by the middle-skill group. Importantly, this information cannot be obtained when data with discontinuities are used instead. We further discuss the possible implications of our finding on the shape of the production function.

Second, we analyze the changes in volatility of hours worked of each occupation group since the mid-1980s, which is relevant to the studies of job polarization in the following sense. Suppose that we want to know if job polarization has made the employment of a specific occupation group more or less stable at the business cycle frequency; that is, a factor that causes job polarization at the low frequency may also cause some occupation group to be disproportionately affected by business cycles. This is equivalent to studying how volatility of the hours variables has changed since the mid-1980s, which is in line with Castro and Coen-Pirani (2008) and Galí and van Rens (2014). We find that the decline in employment volatility for middle-skill occupations is the largest among the three occupational groups. In order to highlight the importance of using consistent data, we report the statistics obtained from inconsistent data. We then analyze the implication of our finding on the welfare cost of business cycles by using a simple partial equilibrium model of consumer's problem. When the change in employment volatility is the only change since the mid-1980s, the welfare cost for the high-skill occupation group after the mid-1980s is about 30 percent lower than before, that for the middle-skill group is about 65 percent lower than before, and that for the low-skill group shows little change.

Last, we show the extent to which different recessionary episodes since the 1980s affect the employment of different skill groups. This is relevant for the studies of business cycles because the group that is affected the most can characterize the recession. For instance, the 1981-82 recession is characterized by substantial drop in employment of the middle-skill occupation group. This is because the manufacturing sector, which heavily depends on the middle-skill group, was severely affected by the recession. Moreover, the 1990-91 recession is associated with a large decline in employment for middle-skill and low-skill occupation groups, while the 2001 recession exhibits the greatest job loss for the low-skill group.

We do not argue that the method of conversion factors is better than the occ1990dd classification system in every respect. As will be discussed in Section 2, we can apply conversion factors when we need "aggregate" hours variables. Instead, the occ1990dd classification system, which constructs the balanced panel structure by occupation, is more useful for micro-level studies,[2] while aggregate hours variables constructed using the method exhibit discontinuities when occupation codes change. Hence, the appropriate method should be considered depending on research subjects.

[2]For example, Autor and Dorn (2013) use the occ1990dd classification to analyze long-term changes in occupation shares based on micro-level data.

2 Data Construction

In this paper, we consider three data sets of aggregate hours variables by occupation. The first data set is the raw data series to which no particular method is applied, and this data set provides a benchmark when evaluating the performance of different methods. The second data set is obtained using the occ1990dd classification that was suggested by Dorn (2009). The last data set, which we mainly use in this paper, is obtained using the conversion factors that were originally developed by the U.S. Census Bureau. While the Bureau of Labor Statistics (BLS) publishes aggregate employment data for different occupation groups that are constructed using conversion factors, there are two major shortcomings associated with using data directly from the BLS. First, its data set covers the period only after 1983. While one can find aggregate employment data for the period before 1983,[3] data for different periods are not directly comparable since the conversion factors are applied only for data after 1983. Second, the BLS publishes employment data only. In this paper, we show that conversion factors can be applied to the period before 1983 to construct consistent aggregate employment data and that the method can also be used to construct a total hours variable, which is not officially provided by the BLS.

The procedure to construct the aggregate hours variables is described in Appendix A.1. In what follows, we introduce data sources and then compare the data set constructed using conversion factors with others in detail.

2.1 Data

There are two data sources used in this paper: the (monthly) Current Population Survey (CPS) Merged Outgoing Rotation Groups (MORG) data that cover the period from January 1979 to December 2010[4]; and the (yearly) March CPS data.[5] Employment data from the March CPS cover the period 1971–2010 while the hours variable in the March CPS covers a shorter period, 1975–2010.

For the purpose of comparison of the different methods, we mainly use the CPS MORG data. In Appendix A.2, we provide additional figures and a table obtained from the March CPS, which share the same properties reported in Section 3.

2.2 Comparison between the occ1990dd Classification System and Conversion Factors

[3]Link: http://fraser.stlouisfed.org/publication/?pid=60.

[4]Data were extracted from the NBER website: http://www.nber.org/data/morg.html.

[5]Data were extracted from the IPUMS website: http://cps.ipums.org/cps. See King, Ruggles, Alexander, Flood, Genadek, Schroeder, Trampe, and Vick (2010).

2.2.1 Overview As is well-known, there have been several main changes in occupation codes in the Census: these have occurred in 1971, 1983, 1992, and 2003.[6] With every introduction of classification systems, new detailed occupations were introduced, some of them were redefined, and some occupations were discontinued.[7] These changes in occupation codes created discontinuities such as substantial drop or jump in employment series between two different occupational classification systems.[8] In particular, the composition of detailed occupations changed substantially in the 2000 system compared to the 1990 system. The main difference between the 1990 and the 2000 systems is the rearrangement of the classification system using the concept of "job families," in which people who work together are classified together, regardless of their skill levels (see Scopp (2003)).[9] As a result of the classification changes, employment data with new occupation codes are not directly comparable with ones for earlier years.[10]

To deal with this inconsistency problem, two methods can be used, although neither is perfect. One is to use the occ1990dd classification system and the other is to use the conversion factors provided by the Census Bureau, which are used to calculate the U.S. official employment data by occupation.

2.2.2 occ1990dd Classification System The occ1990dd classification system provides a balanced panel of occupation codes covering the 1980, 1990, and 2000 Censuses and the 2005 ACS.[11] This system provides a one-to-one mapping of Census occupation codes to a unified category system. As a result, a single occupation code in a specific Census scheme corresponds to another single occupation code in a unified category system (occ1990dd).

The IPUMS-CPS website provides occupation codes under the "occ1990" system, which makes occupations in each Census year comparable to the 1990 Census, based on Meyer and Osborne (2005).

[6]The Census introduces a new classification of occupations every 10 years. The CPS uses the 1970 occupation codes from 1971 to 1982, the 1980 occupation codes from 1983 to 1991, the 1990 occupation codes from 1992 to 2002, and the 2000 occupation codes starting from 2003. For example, in January 2003, the CPS adopted the 2000 occupation codes replacing the 1990 Census occupational classification system. Due to the code changes, direct comparison between the estimates from different classification schemes is not possible without major adjustments.

[7]Even when the names of occupations appear to be the same, the definitions of the categories are sometimes different. For example, "Managers" and "Farming Occupations" were defined differently in 1990 from the way they were defined in 2000, even though these words appear in occupation tables from both Censuses (Scopp (2003)).

[8]The IT and the general economic boom in the U.S., observed particularly in the late 1990s, and the peaking of oil prices during the early-to-mid 2000s would have contributed to some structural breaks in employment, but in this paper, we focus on breaks which occur due to changes in classification codes.

[9]For example, farmers and farm managers are now under the same group with management occupations.

[10]While there was a significant change in the Census occupation coding scheme between the 1970 and 1980 Censuses and between the 1990 and 2000 Censuses, there were relatively few changes in the Census occupation coding scheme between the 1980 and 1990 Censuses, allowing for suitable comparisons over this period.

[11]Some parts of the discussion in this section are from Dorn (2009).

Meyer and Osborne (2005) start from a weighted crosswalk and construct an unweighted (and thus easier to use) crosswalk that approximates the weighted correspondence. Meyer and Osborne (2005) assign an occupation code from the 1990 Census to each individual in the 2000 Census using the most likely match according to the Census Bureau's official crosswalks. In several cases, however, the assigned codes are misallocated (Lefter and Sand (2011)). In addition, the occ1990 classification system does not provide a "balanced" panel of occupations (Dorn (2009)). For instance, "Economics Instructor" was included in the occ1990 system, but there are no observations for this occupation in the 2000 Census because the specific fields of college teachers were not reported in the 2000 Census. Dorn (2009) shows that this unbalanced structure can be problematic for empirical work related to employment or wage changes within detailed occupations. Most of the improvements in the occ1990dd classification system were made by making occupation definitions "broader" (by simply aggregating across occ1990 occupations) and providing more "consistent" definitions in service occupations. That is, the occ1990dd system is primarily an aggregation of occ1990 codes in order to create time-consistent occupation categories and a balanced panel of occupation codes.

However, there is still a possible drawback associated with using the occ1990dd classification system. Since the occ1990dd code provides a one-to-one correspondence between occupation titles in different Census schemes, it may not capture the fact that not every worker in a Census category falls into another single Census category. As shown in the next section, there exists a proportional flow between different Census categories. Therefore, the occ1990dd classification system may provide a biased estimate of labor market variables, especially when studying short-term fluctuations of aggregate occupational employment. Although the occ1990dd is not intended to be used for analyses of short-run changes in the labor market, several papers (for instance, Foote and Ryan (2014) and Gaggl and Kaufmann (2015)) have employed this method to study business cycle properties of occupational employment. They corrected discontinuities by leveling up or down as a temporary expedient. Instead, we introduce a more convenient way of addressing the inconsistency issue in the next section.

2.2.3 Method of Conversion Factors

The conversion factors are weighted crosswalks that show the proportional flows for individual occupation categories between the two Census years in order to bridge changes in occupation codes.[12] As illustrated by Blau and Liu (2013), this is important because

[12] If researchers are interested in doing an analysis for a short time period in which there is one consistent set of Census occupation codes (for instance, 1992–1999), they should use the original code, occ1990, provided by the IPUMS, not

all the incumbents of a particular Census occupation group do not necessarily match with the other Census occupation group, but instead are split into several categories.

For a particular occupation within each Census, the Census crosswalks show lists of occupations in other Censuses. Since there is not always a one-to-one match between the Census classifications, we need to convert one scheme to the others. The Census Bureau provides conversion factors to help data users bridge the gap created by discontinuities in occupational series. The conversion factors show the percentage distribution of employment from each occupation code in one classification, for example, 1990, into each code in the other classification, say, 2000. The Census Bureau provides crosswalk tables and conversion factors between the 1990 and 2000 Census classifications. These factors are based on three-year average survey microdata (2000–2002, double-coded sample) that were coded to both the old and new classification systems. This process puts each person in the sample into both classification sets.[13] The conversion factors, thereby, provide information on the proportion of workers that went from one Census category into another.

For example, among "Executive, administrative, and managerial" workers in the 1990 occupation group, the Census crosswalk distributes 73.1 percent of incumbents to "Management, business, and financial operations occupations" in the 2002 occupational group, 11.1 percent to "Office and administrative support occupations," 4.4 percent to "Sales and related occupations," 4.1 percent to "Professional and related occupations," 4 percent to "Service occupations," and 2.2 percent to "Construction and production occupations."[14] As a result, the conversion factors provide a crosswalk of the most likely occupation that a worker coded under one classification scheme would have been coded under another scheme.[15]

Although the conversion factors provide linkages between the old and new classifications, there are

occ1990dd or conversion factors. If their analysis spans multiple time periods, however, they have to work with consistent crosswalks as the Census has changed the occupation codes every 10 years.

[13]The Census Bureau uses a similar method to create conversion factors for earlier Censuses with the 1970 Census sample comprising 127,125 individuals drawn from the 1970 Census (Mosbacher and Ortner (1989)). This sample was used to determine what fraction of cases with a given 1970 code correspond to each of several 1980 codes. For converting 1980 to 1990 classifications, only a few changes were needed to make the 1980 codes compatible with the 1990 classifications as there were relatively few changes in the Census occupational coding scheme between the 1980 and 1990 Censuses.

[14]The Census Bureau also created conversion factors for industry employment. For example, for the construction industry, 92 percent of employment in the old construction classification was re-classified to the new construction classification. The remaining 8 percent of employment was re-classified among other industry categories in the new classification (link: http://www.bls.gov/cps/constio198399.htm).

[15]In theory, it is also possible to convert the 2000 categories into those for 1990 in order to make reverse comparisons over time (see Scopp (2003)). However, the Census Bureau recommends using the Census crosswalk only for converting the codes forward to display employment of 1990 in 2000, as the 2000 Census is more up-to-date. Our results are not much affected by converting direction.

undoubtedly some limitations in how they can be used. Scopp (2003) reports that conversion factors are subject to sampling errors, especially when the numbers for a detailed category are very small. Thus, the resulting series is only an approximation providing a general employment trend over time.[16] Further, the double-coding process might involve coding errors. These errors may contaminate comparisons across classifications. In addition, while conversion factors allow us to analyze employment and total hours worked by detailed occupations, this method is not appropriate for other variables such as average hours worked and real wages by occupation. For instance, if we apply conversion factors to construct the hourly wage rate by each occupation group, we can observe breaks between 2002 and 2003, when there was a change in occupation codes. Nevertheless, the conversion factor method has a few advantages compared to other methods, which will be shown in the following section.

3 Comparing Consistency of Different Occupational Data

In this section, we present figures for the aggregate hours variables (i.e., employment and total hours) to show that aggregate data constructed by conversion factors outperform other data sets in terms of consistency.[17]

3.1 Employment We first consider employment in Figures 3.1 to 3.3. In each figure, the solid blue line is the employment series constructed using conversion factors, the thick green dotted line is constructed using the occ1990dd classification, and the thin red dotted line is from the raw data. The shaded regions are the official NBER recession dates.

We can observe that the employment series constructed using the occ1990dd classification have three major breaks over time. Employment of high-skill occupations exhibits two breaks between 1982 and 1983 and between 2002 and 2003. The other break is observed in Figure 3.3: employment of low-skill occupations shows a break between 2002 and 2003. The downward break for employment of high-skill and the upward break for that of low-skill occupations between 2002 and 2003 are related to the concept of job families, which was introduced in the 2000 Census occupational classification system (see Section 2.2.1). For example, the 2000 Census system places first-line supervisors in the same group as the workers they supervise in the manual group, and thus using the occ1990dd classification artificially increases the

[16] Link: http://www.bls.gov/cps/cpsoccind.htm.

[17] To download consistent aggregate data, visit the authors' websites.

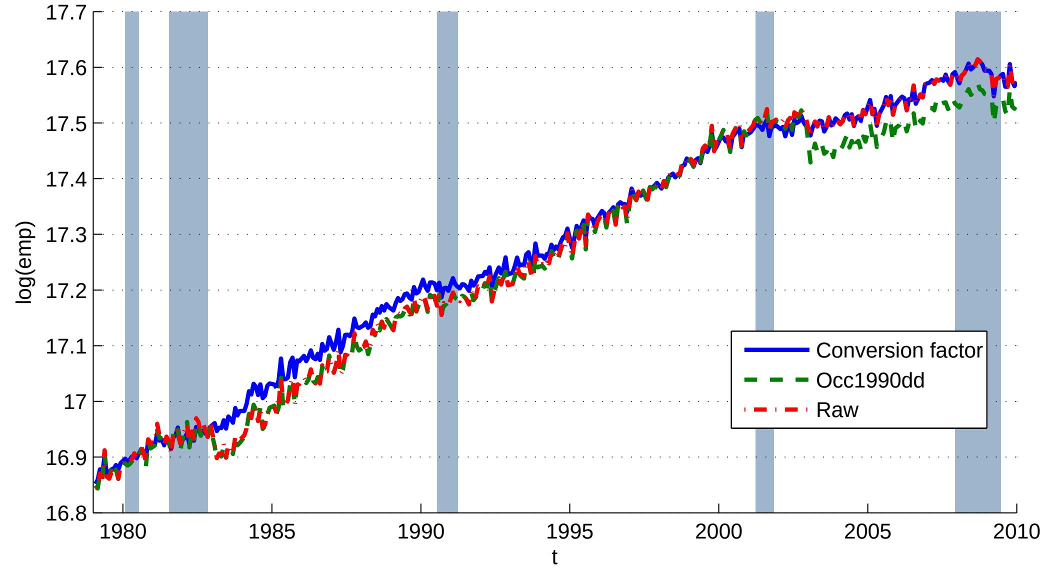

Figure 3.1: Employment: High-Skill Occupations

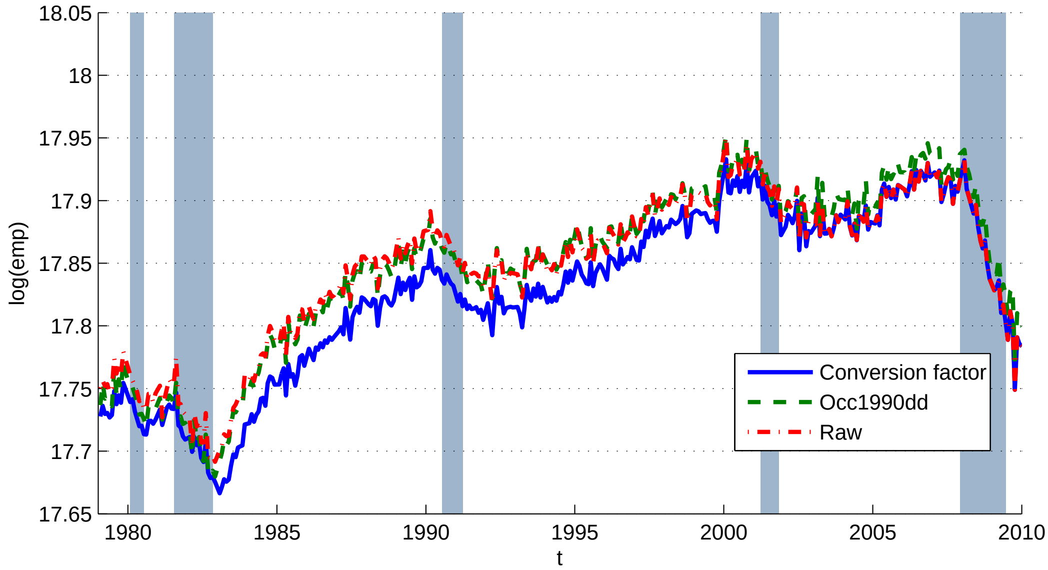

Figure 3.2: Employment: Middle-Skill Occupations

employment of low-skill occupations. In addition, the level of employment for high-skill occupations is lower for data constructed using the occ1990dd classification than for data using conversion factors since 1983, while the level of employment for middle-skill occupations constructed using conversion factors is lower than the other series using the occ1990dd classification.

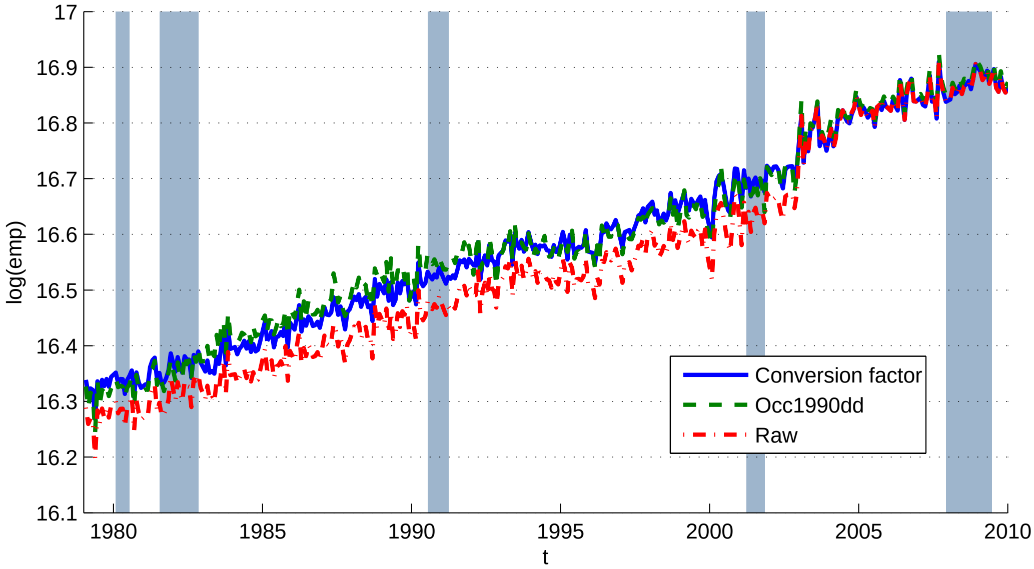

Figure 3.3: Employment: Low-Skill Occupations

Similarly, the raw data exhibit two breaks: one for high-skill occupations between 1982 and 1983 and the other for low-skill occupations between 2002 and 2003.[18] This implies that the occ1990dd classification system does not resolve the inconsistency problem associated with using aggregate employment data; rather, it amplifies the inconsistency problem. In contrast to these data sets, the data we construct using conversion factors exhibit no discontinuities in any of the series; in other words, the conversion factor method yields the most consistent employment series among the three methods. Applications of the above findings, which are presented in Section 4, will make it clear why having a more consistent data set is important in studies of business cycle properties of job polarization.

3.2 Total Hours Worked

Now we move our attention to total hours worked for each occupation group. Figures 3.4 to 3.6 are the corresponding figures. One can easily observe that the total hours

[18] A Chow test for those breaks in the occ1990dd data and the raw data rejects the null hypothesis of no break at the 5 percent level and confirms that there are structural breaks as shown in figures above. If we look at the magnitude of those breaks in terms of employment growth rates, it is apparent for the series to have discontinuities. For instance, between 1982 December and 1983 January, the data from conversion factors show that the employment of high-skill occupations changes by less than 1 percent, while the occ1990dd data and the raw data show that the high-skill employment decreases by 3.5 percent and 4.2 percent, respectively.

However, the data constructed from conversion factors also exhibit breaks when there are legitimate reasons such as business cycles. For example, between 2002 and 2003, all the data show jumps in low-skill occupations due to the recovery following the early 2000s recession. However, the magnitude in the data from conversion factors is smaller than others: the data from conversion factors show that the low-skill employment increases by 9.2 percent, while the occ1990dd data and the raw data exhibit 13.1 percent and 14.1 percent jumps, respectively.

worked series constructed using the conversion factors does not exhibit any discontinuities, which confirms that the conversion factor method yields a more consistent total hours series than the occ1990dd classification system.

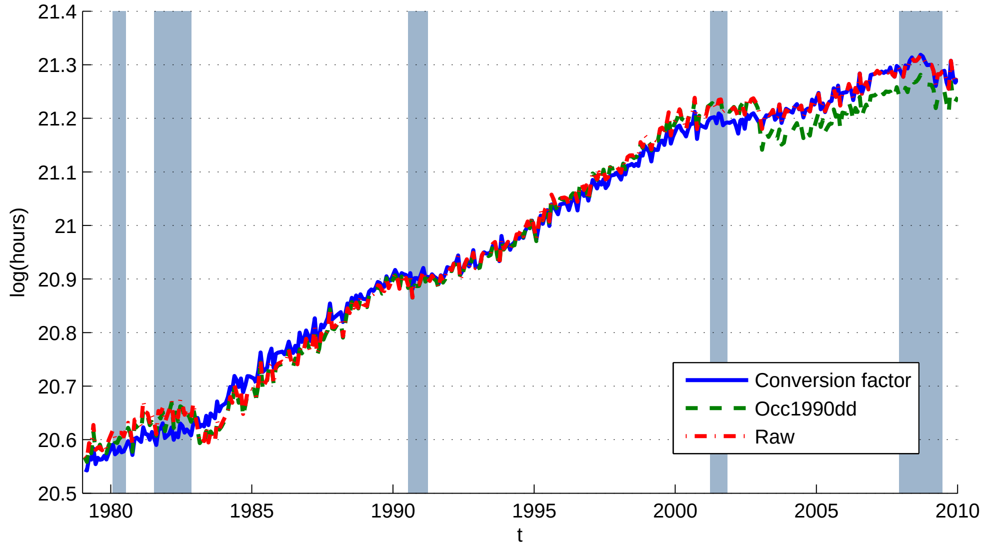

Figure 3.4: Total Hours: High-Skill Occupations

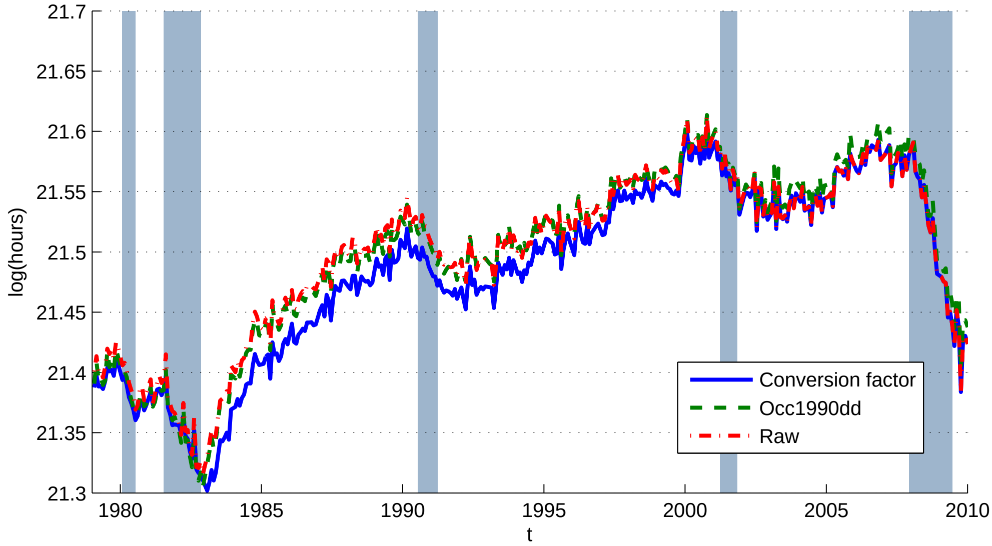

Figure 3.5: Total Hours: Middle-Skill Occupations

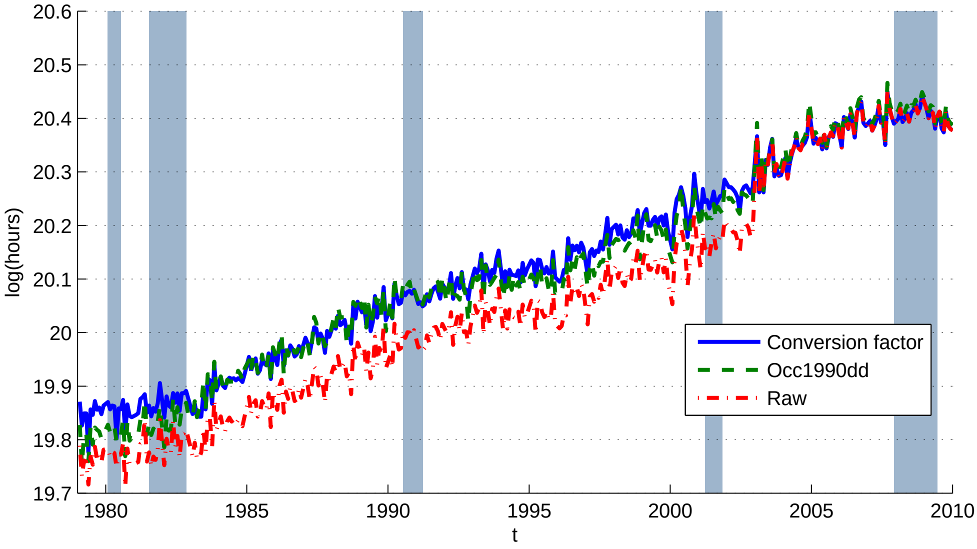

Figure 3.6: Total Hours: Low-Skill Occupations

4 New Stylized Facts on Business Cycle Properties of Occupational Employment

In this section, using our consistent data, we present several new empirical insights on business cycle properties of occupational employment. In order to emphasize the importance of using consistent data, we first consider a researcher who is interested in the long-run changes in employment. Then, the discontinuity we observe from the pre-existing data may not be a problem because they do not alter information on long-run trends of employment. For instance, if a researcher is only interested in the occurrence of job polarization, the disappearance of jobs for the middle-skill group during the last 30 years is still observed even when data are inconsistent.[19] However, observations on business cycle properties of job polarization change dramatically when we use inconsistent occupational data instead of consistent data. Furthermore, the major changes in the occupation codes in 1983, 1992, and 2003 nearly coincide with the three recessions whose NBER official dates are July 1981–November 1982, July 1990–March 1991, and March 2001–November 2001, respectively. Hence, inconsistent data with discontinuities in these periods would provide us with inaccurate information on the group that was mostly affected by these recessions. In this section, we compare the findings obtained from our data and

[19]Although the growth rate of employment may differ across different data sets, it does not hide job polarization.

those from inconsistent data whenever possible in order to highlight the importance of using consistent employment data.

For the analysis below, we remove the seasonality in the data by using X-12 ARIMA and then using the Baxter and King (1999) filter with $\kappa = 12$ (size of the window for the filter) when we detrend the data. As usual, we set the lowest frequency at 6 quarters and the highest frequency at 32 quarters.[20]

4.1 Revisiting the Technology-Hours Debate: Heterogeneous Responses by Occupation

Since the influential work by Galí (1999), whether a technology shock increases or dampens hours (or employment) has been a controversial issue in the literature (see Francis and Ramey (2005), Chang and Hong (2006), Francis and Ramey (2009), and Caporalea and Gil-Alana (2014) among others) because it provides a direct empirical test on the validity of the existing business cycle theories. However, possible heterogeneity in the effect of a technology shock on different occupations has been widely neglected; while Balleer and van Rens (2013) study the effect of (investment-specific) technology shocks on relative hours between skilled and unskilled workers, they only consider two groups of workers in the analysis and focus more on skill-biased technology changes[21] rather than the neutral technology shocks.

In this section, we fill this void in the literature. In doing so, we estimate the VAR system, following Galí (1999)[22]:

$$\begin{bmatrix} \Delta x_t \\ \Delta n_{it} \end{bmatrix} = \begin{bmatrix} C^{11}(L) & C^{12}(L) \\ C^{21}(L) & C^{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_t^z \\ \varepsilon_t^{nz} \end{bmatrix} \equiv C(L)\varepsilon_t \tag{4.1}$$

where $\Delta x_t \equiv x_t - x_{t-1}$ where x_t is the log labor productivity, $\Delta n_t \equiv n_{it} - n_{it-1}$ where n_{it} is the log employment of occupation group $i \in \{\text{high-skill, middle-skill, low-skill}\}$, L is the lag operator, $\varepsilon_t = [\varepsilon_t^z, \varepsilon_t^{nz}]'$, $\mathbb{E}\varepsilon_t\varepsilon_t' = I$ and ε_t^z (resp. ε_t^{nz}) is the technology shock (resp. non-technology shock). Labor productivity is measured in the non-farm business sectors and is downloaded from the BLS website. Since the data on labor productivity are at quarterly frequency, we take the average of three months to obtain the quarterly employment series for each occupation group. Consistent with Galí (1999), the standard augmented Dickey-Fuller (henceforth ADF) test does not reject the null of a unit

[20]Our results do not change even when we use the Hodrick–Prescott filter.
[21]See Mallick and Sousa (2016) for related discussions on skill-biased technology changes.
[22]For the estimation, we use the VAR code shared by Cesa-Bianchi (2015).

root in the levels of each series but rejects the same null when it is applied to the first-differences (at the 1 percent level) so that we use the log difference in the estimation.[23]

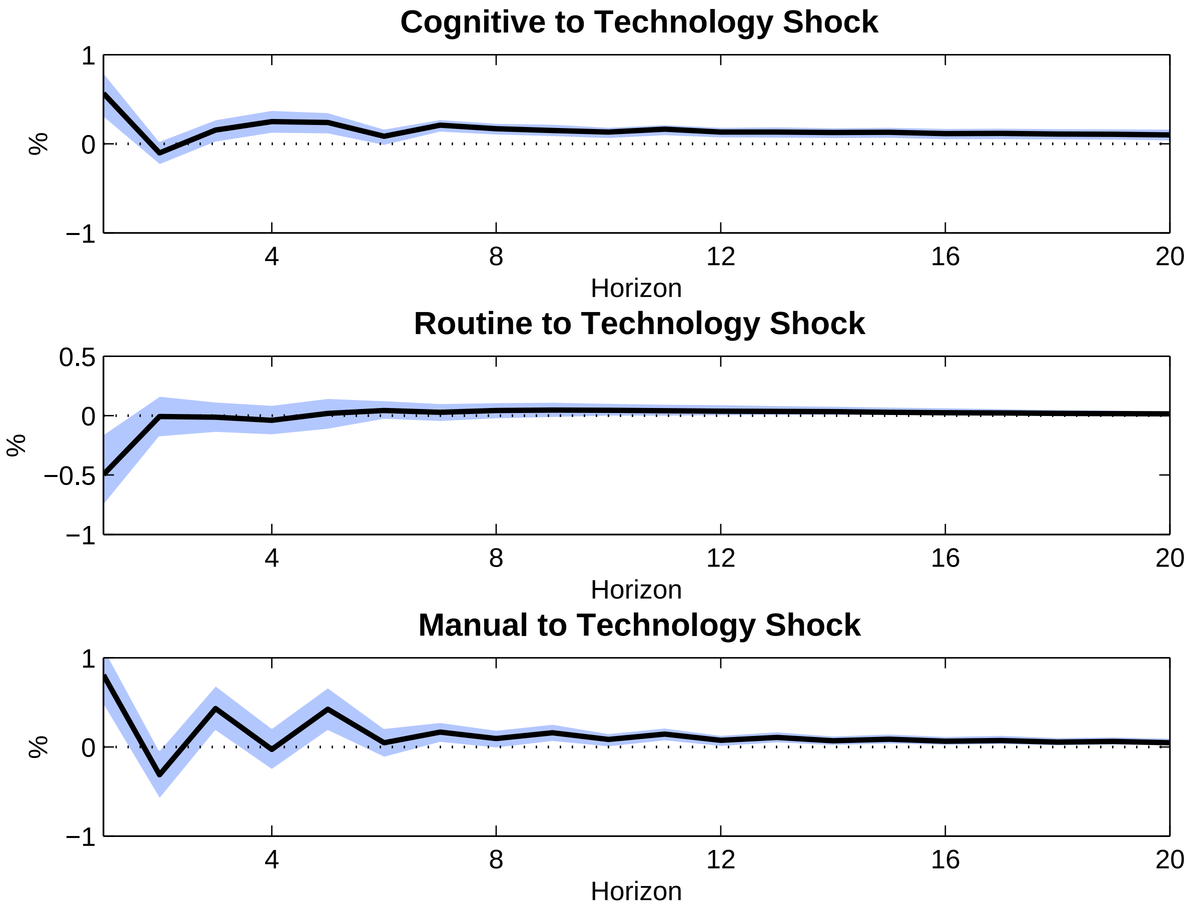

Figure 4.1: Impulse Response Functions: Technology Shocks
Note: Each graph displays the IRFs with bootstrapped 90 percent confidence intervals. Impulse responses are multiplied by 100 and are denoted by %.

In order to identify the technology shock, we use the long-run restriction à-la Blanchard and Quah (1989) and Galí (1999), which is the most commonly adopted identification strategy in the literature: $C^{12}(L) = 0$, and hence, we assume that it is the technology shock that has long-run effects on labor productivity.[24] Figure 4.1 shows the impulse response functions of each employment with respect to

[23]See also Cover and Mallick (2012) for related discussions. While their shock identification strategy is to impose the short-run restriction instead of the long-run restriction usually imposed in the literature, the impulse response of the unemployment rate to the technology shock shows similar patterns.

[24]One advantage of the long-run restriction is that the ordering of variables in estimation is not critical while it may critically affect estimation results when the short-run restriction such as Cholesky decomposition is adopted for identifi-

the technology shock.[25] The 90 percent confidence intervals are computed by a bootstrap method. Interestingly, the employment of high-skill and low-skill occupation groups increases when there is a positive technology shock. The employment of the middle-skill group, on the contrary, exhibits the opposite pattern; it declines after the technology shock hits the economy. This finding implies that the key finding of Galí (1999) and Francis and Ramey (2005), that hours drop upon the arrival of the technology shock, can be mainly driven by the response of the middle-skill group to the technology shock. Our observation can be summarized as follows:

Stylized Fact 1 (Employment Responses to Technology Shock)**.** *During the sample period 1979–2010, the middle-skill occupation group is negatively correlated with the technology shock. The high-skill and low-skill groups react positively to the technology shock.*

This finding suggests that it is important to identify the channels through which the technology shock affects the aggregate labor market. Among many possible channels, we briefly discuss the "capital-labor substitution" channel, which also provides an insight on the shape of the production function. In the job polarization literature (for instance, Autor, Levy, and Murnane (2003)), it is usually argued that the middle-skill occupation is a relative substitute for capital while the high-skill and low-skill occupation groups are relative complements to capital. This is consistent with the above empirical fact in the following sense. When the productivity shock hits the economy, capital accumulates over time. Then, the employments of high-skill and low-skill groups also increase as they are complements to capital. However, the middle-skill employment decreases because it is a relative substitute for capital. Suppose instead that all occupation groups are equally complements to capital; that is, the production function takes the Leontief form. Then, the impulse responses would take the same sign among the three groups, when the shock hits the economy because of the complementarity, unless the price rigidity affects different occupation groups disproportionately, which is inconsistent with Figure 4.1.

4.2 Changes in Volatility of Occupational Employment

In this subsection, we analyze which jobs are less volatile than others at the business cycle frequency and whether there have been any changes in that information since the mid-1980s. The former can be seen as a question of seeking

cation. Moreover, the long-run restriction is the conventional way to incorporate one of the key ideas of the real business cycle theory into the shock identification that the technology shock is important in determining labor productivity. See Blanchard and Quah (1989) and Galí (1999) for more detailed discussions.

[25] While not reported, employment unambiguously increases for all groups when the non-technology shock hits the economy, which is consistent with Galí (1999).

jobs for which employment fluctuates less than others during business cycles. The second question is noteworthy for the following reasons. Job polarization, which shifts firms' demand from the middle-skill group to the high-skill and low-skill groups, started to occur in the mid-1980s. It is possible that such job polarization also affects the volatility of jobs in an asymmetric way. For instance, the middle-skill occupation group, whose importance in production has decreased over time, may suffer more from business cycle fluctuations than before, while other groups do not. Hence, a study of the changing cyclical properties of employment deepens our understanding of job polarization. Moreover, given the changes in the aggregate economy since the mid-1980s, such as the vanishing procyclicality of labor productivity (see Galí and van Rens (2014)) and the increase in wage volatility relative to GDP (see Champagne and Kurmann (2013)), this study further enhances our knowledge of the labor market.

We compute the standard deviation of each detrended employment series by dividing the whole period into two: (1) 1979–1983 and (2) 1984–2010, following Castro and Coen-Pirani (2008) and Galí and van Rens (2014).[26] One might raise a concern about the validity of the statics from the first subperiod because it only covers five years. As a robustness check, we do the same exercise with the March CPS, whose employment series cover the period from 1971 to 2010, so that the first subperiod is longer; the result is reported in Table A.1.[27] We find that the results from the two data sets are similar.

Table 4.1 shows the main result and Stylized Fact 2 summarizes the key observations from the table.

Table 4.1: Standard Deviations of Detrended Employment

Occupation Group	High-Skill		Middle-Skill		Low-Skill	
Period	(1) 79–83	(2) 84–10	(1) 79–83	(2) 84–10	(1) 79–83	(2) 84–10
Standard Deviation	0.39	0.33 (0.83)	0.84	0.50 (0.60)	0.51	0.52 (1.04)

Note: 1. All numbers are multiplied by 100.
2. Numbers in parentheses are the ratios between (2) and (1).

Stylized Fact 2 (Volatility Changes)**.** *Let $\Delta\sigma_i$ be the percentage change in employment volatility of occupation group i. Then, $\Delta\sigma_M < \Delta\sigma_H < 0 < \Delta\sigma_L$. In addition, in absolute terms, the change in the volatility of detrended employment since the mid-1980s is the greatest for the middle-skill occupation group.*

A key finding from the consistent data is that employment volatility decreases the most for middle-

[26] Data from 1979 and 2010 are truncated as we set $\kappa = 12$.
[27] For the March CPS, we set 2 as the lowest and 8 as the highest frequency with $\kappa = 3$.

skill occupations. This finding is important because the labor market has changed unfavorably for the middle-skilled workers in the long-run since the mid-1980s (i.e., job polarization) while the changes in employment volatility at the business cycle frequency have become more favorable for them relative to other occupation groups. For the low-skill occupation group, there has been virtually no change in employment volatility since the mid-1980s. As a result, the employment volatility of the middle-skill occupation group was about 2.1 (resp. 1.6) times higher than that of the high-skill (resp. low-skill) group before 1984 but it is about 1.5 times higher (resp. almost same) in the second sub-period. This finding provides detailed information than one can learn from Castro and Coen-Pirani (2008): After classifying workers into two skill groups according to educational attainments, where the unskilled group includes middle-skilled and low-skilled workers under our classification, they show that employment volatility drops for the unskilled group. Stylized Fact 2 unveils the fact that while information on occupation is used instead of education in our analysis, the drop in employment comes from middle-skilled workers, not from low-skilled workers.

What would happen if we use data with discontinuities? Table 4.2 provides the results with inconsistent data and Stylized Fact 3 summarizes the findings.

Table 4.2: Standard Deviations of Detrended Employment: Inconsistent Data

Occupation Group	High-Skill		Middle-Skill		Low-Skill	
Methodology\Period	(1) 79–83	(2) 84–10	(1) 79–83	(2) 84–10	(1) 79–83	(2) 84–10
occ1990dd System	1.09	0.48 (0.45)	0.88	0.52 (0.59)	0.47	0.60 (1.27)
Raw Data	1.11	0.37 (0.44)	0.87	0.52 (0.59)	0.56	0.73 (1.31)

Note: 1. All numbers are multiplied by 100.
2. Numbers in parentheses are the ratios between (2) and (1).

Stylized Fact 3 (Volatility Changes: occ1990dd data and Raw Data)**.** *Let $\Delta\sigma_i$ be the percentage change in employment volatility of occupation group i. Then, $\Delta\sigma_H < \Delta\sigma_M < 0 < \Delta\sigma_L$ when inconsistent data are used: employment volatility decreases significantly for high-skill and middle-skill occupations, while it increases significantly for low-skill occupations.*

When the occ1990dd data or raw data are used, the decline in the standard deviations of employment is larger for high-skill occupations than for middle-skill occupations: the standard deviation for high-skill occupations decreases to less than half of its previous level. We note that the consistent data set shows a smaller decline for the high-skill group: the employment volatility for them decreases by less than

20 percent. Further, the differences are dramatic when the low-skill occupation group is considered. Employment data constructed using conversion factors show that there has been virtually no change in volatility for the low-skill group. However, this increases by about 30 percent when using the occ1990dd data or raw data. In contrast to these two groups, there is no substantial difference for the middle-skill occupation group because none of the three methods exhibits artificial breaks for this group. Further analysis on why the statistics are different across data sets is provided in Appendix A.3.

In order to show the importance of our finding, we provide one implication on studies of business cycles. As is well-known, the welfare cost of business cycles can be easily approximated by the volatility of consumption (Lucas (1987)); namely, $\lambda \propto \mathbb{V}(c_t)$, where λ measures the extent to which a consumer would be willing to pay to avoid fluctuations so that it measures the welfare cost of business cycles and $\mathbb{V}$ represents the variance of consumption series. Then, one can prove the following.

Proposition 1 (Welfare Cost of Business Cycles)**.** *Consider the following consumer utility maximization problem.*

$$\max_{\{c_t, a_{t+1}\}_{t=0}^{\infty}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \ln c_t \right] \tag{4.2}$$

subject to

$$(1)\ c_t + a_{t+1} = wh_t + (1+r)a_t$$

$$(2)\ a_0 = a \geq 0 \textit{ is given}$$

where w is the wage rate, a_t is the asset holding, a is the steady state wealth level, and h_t represents the hours worked of the consumer.

If (1) hours worked (employment) are exogenously given as an AR (1) process ($\hat{h}_t = \rho_h \hat{h}_{t-1} + \varepsilon_t$, where $\hat{x} \equiv \ln(h_t/h)$, h is the steady state hours, $\rho_h \in (0,1)$, and $\varepsilon_t \sim N(0, \sigma_h^2)$), (2) the asset market is incomplete, (3) a consumer takes prices as given, and (4) wage rates and skill premia are acyclical, then the welfare cost of business cycles can be approximated as follows.[28]

$$\lambda \propto \sigma_h^2 \tag{4.3}$$

[28] Results on the acyclical property of wage rates and skill premia are available upon request.

Proof. See Appendix A.4. □

Therefore, we can simply compute the contribution of changes in hours volatility to the welfare cost of business cycles under the assumption that other variables and parameters do not change significantly over time. From Table 4.1, we can compute σ_h^2 as $\sigma_h^2 = (1-\rho_h)\mathbb{V}(h_t)$, where $1-\rho_h \approx 100$, so that $\sigma_h^2 = 100\mathbb{V}(h_t)$. Then, the welfare cost of business cycles for the high-skill occupation group after the mid-1980s is about 30 percent lower than before, that for the middle-skill group is about 65 percent lower than before, and that for the low-skill group remains nearly unchanged in our simple framework. Therefore, the changes in employment volatility since the mid-1980s have been more favorable for middle-skilled workers when other changes are isolated. This finding has a further implication on the labor market policy: If a government cares about the welfare of workers with different skills, a policy that focuses on the middle-skill occupation group would have become less important since the mid-1980s at the business cycle frequency.

However, if we instead use the inconsistent data constructed through the occ1990dd classification system, the implication is totally different for high-skilled and low-skilled workers; the welfare cost for the high-skill occupation group decreases by 80 percent after the mid-1980s and that for the low-skill occupation group increases by 60 percent. Hence, when we use the information obtained from inconsistent data, the changes in employment volatility at the business cycle frequency would be more favorable for high-skilled workers than for middle-skilled workers.

4.3 Asymmetric Effects of Recessions on Different Occupation Groups

In this subsection, we consider the heterogenous effects of recessions on different occupations and show that the results depend critically on which data we use. The question on the effects of recessions is important in terms of defining the characteristics of a recession. For example, the 2007 recession originated from the collapse of the housing market, and hence, the construction and financial sectors were mostly affected by the recession. Meanwhile, the health care sector expanded during the recession (Sahin, Song, and Hobijn (2010)). As Jaimovich and Siu (2014) have found, the 2007 recession affected each occupation disproportionately and deepened job polarization.

However, inconsistent data sets do not allow careful analysis in this regard: recall that the major code changes in occupations occurred in 1983, 1992, and 2003, and these years nearly coincide with three recessions after 1980 (i.e., July 1981–November 1982, July 1990–March 1991, and March 2001–November

2001). Hence, inconsistent data with discontinuities in these periods provide inaccurate information on which occupational group was affected most severely by recessions. Our consistent data, however, make it possible to analyze the asymmetric effect of recessions.

For this exercise, we compare the detrended employment series constructed using conversion factors with the series obtained by applying occ1990dd classification.[29] Figures 4.2 to 4.4 show the results.[30] The shaded region corresponds to the official NBER recession dates, the solid blue line is employment of the high-skill occupation group, the thick dotted green line is employment of the middle-skill occupation group, and the thin dotted red line is employment of the low-skill occupation group.

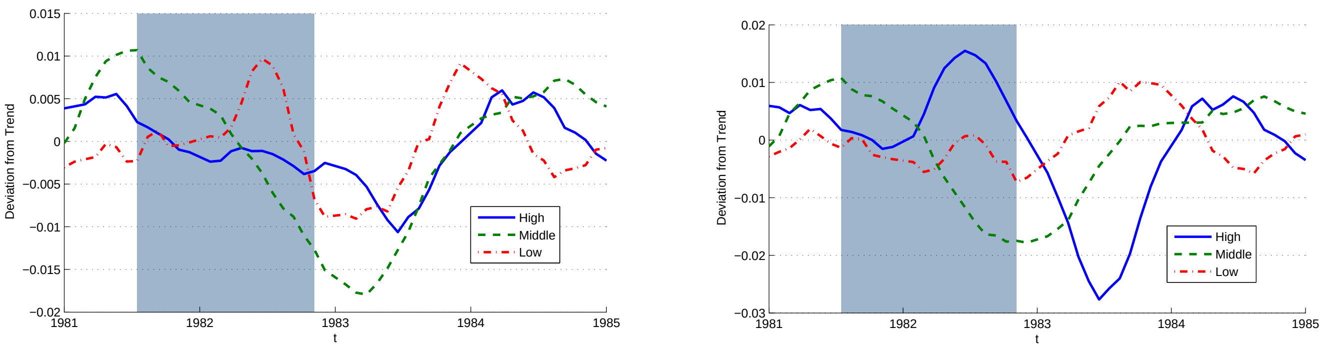

Figure 4.2a: Method of Conversion Factors

Figure 4.2b: occ1990dd Classification System

Figure 4.2: Employment Fluctuations During the 1981-82 Recession

We first consider the 1981–82 recession. Figure 4.2a, constructed using conversion factors, shows that the group that loses jobs the most is the middle-skill occupation group, while the high-skill occupation group is not significantly affected by the recession. However, Figure 4.2b, derived from the occ1990dd system, shows that employment fluctuations of high-skill occupations are greater than those of middle-skill occupations. The effect of the 1981–82 recession was asymmetric across sectors; in particular, the manufacturing sector, which depends heavily on the middle-skill occupation group, was affected the most by the recession. Accordingly, Figure 4.2b leads to a misunderstanding of the differences in employment fluctuations across occupation groups during the 1981–82 recession.

Similar to the 1981–82 recession, it is evident that the different implications for the 1990–91 recession are derived from different data sets. Figure 4.3b shows that employment of low-skill occupations declines by more than 1 percent after the end of the recession. This is slightly greater than the decline in employment for middle-skill occupations (less than 1 percent); hence, it may seem that the

[29] The analysis below using detrended data is not conducted by Foote and Ryan (2014) because of the problem of data inconsistency. Hence, discussions in this section supplement their analysis.

[30] We get similar results when we use the total hours series.

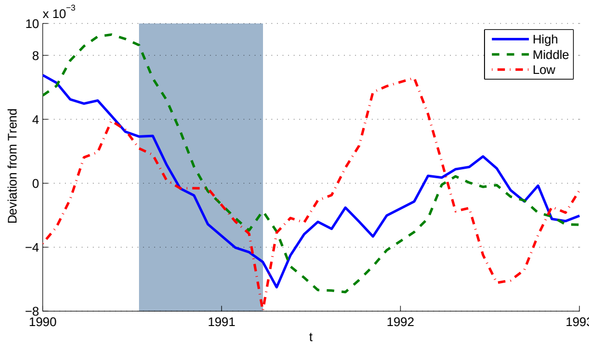

Figure 4.3a: Method of Conversion Factors

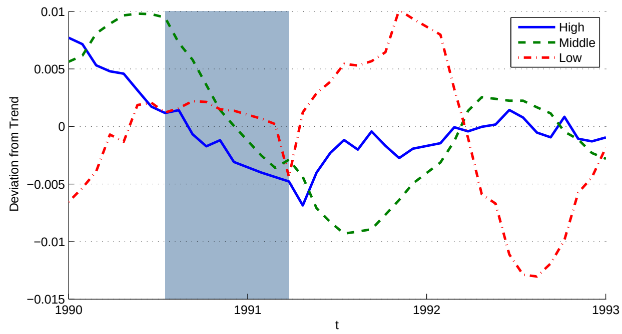

Figure 4.3b: occ1990dd Classification System

Figure 4.3: Employment Fluctuations During the 1990–91 Recession

1990–91 recession affected low-skill occupations most severely in terms of employment fluctuations. However, employment for the low-skill occupation group declines by about 0.8 percent when consistent employment data are instead used, which is similar to that for middle-skill occupations.

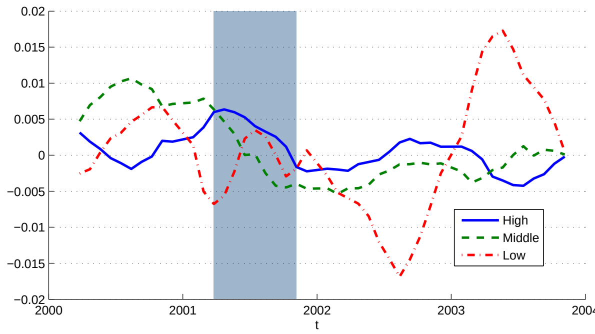

Figure 4.4a: Method of Conversion Factors

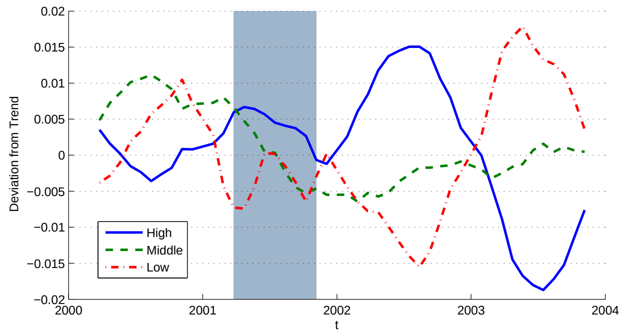

Figure 4.4b: occ1990dd Classification System

Figure 4.4: Employment Fluctuations During and After the 2001 Recession

Figure 4.4 plots the employment fluctuations of each occupation group for another recessionary episode, that is, the 2001 recession. One can easily observe that the two graphs in the figures show different patterns of recovery from the recession. Figure 4.4a, derived from consistent data, shows that employment of the high-skill occupation group changed little during the episode. However, Figure 4.4b, constructed using the occ1990dd classification, shows that the high-skill occupation group exhibits more than a 3 percent drop in employment during the recovery, suggesting the misleading conclusion that there was a lagged response of the high-skill occupation group to the recession.

5 Conclusion

We have discussed the value of having consistent aggregate employment data through the method of conversion factors in studies of business cycle properties. In terms of methodology, we show that (1) the method of conversion factors can also be used to construct total hours worked; (2) it can be applied to the period before 1983 where such data are not provided by the BLS; and (3) employment and total hours worked, which are constructed using conversion factors, are more consistent than previously available data sets are. Equipped with consistent data, we study three aspects of occupational employment fluctuations over the business cycles and discuss their implications on studies of business cycles. In particular, we investigate (1) how the occupational employment responds to technology shocks, (2) the changes in employment volatility of each occupation group since the mid-1980s, and (3) the extent to which the previous recessions affect occupation groups differently.

We finally note that while the occ1990dd classification system is not successful in constructing consistent aggregate data, it is useful in micro-level studies that require unambiguous assignment of individual workers to specific occupation codes. By contrast, the method of conversion factors cannot be used in micro-level studies since it is only useful for constructing aggregate data instead of tracking each individual. Hence, researchers should be cognizant of the different advantages of using different methodologies when analyzing occupational employment.

References

Acemoglu, D., and D. H. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 4, 1043–1171.

Autor, D. H., and D. Dorn (2013): "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market," *American Economic Review*, 103(5), 1553–1597.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?," *Quarterly Journal of Economics*, 113(4), 1169–1213.

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

Autor, H. D. (2010): "The Polarization of Job Opportunities in the U.S. Labor Market: Implications for Employment and Earnings," *The Hamilton Project*.

Balleer, A., and T. van Rens (2013): "Skill-Biased Technological Change and the Business Cycle," *Review of Economics and Statistics*, 95(4), 1222–1237.

Baxter, M., and R. G. King (1999): "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series," *Review of Economics and Statistics*, 81(4), 575–953.

Beaudry, P., D. A. Green, and B. M. Sand (2015): "The Great Reversal In the Demand for Skill and Cognitive Tasks," *Journal of Labor Economics*, Forthcoming.

Blanchard, O. J., and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79(4), 654–673.

Blau, Francine D., P. B., and A. Y.-H. Liu (2013): "Trends in Occupational Segregation by Gender 1970-2009: Adjusting for the Impact of Changes in the Occupational Coding," *Demography*, 50(2), 471–492.

Caporalea, G. M., and L. A. Gil-Alana (2014): "Persistence and cycles in US hours worked," *Economic Modelling*, 38, 504–511.

Castro, R., and D. Coen-Pirani (2008): "Why Have Aggregate Skilled Hours Become So Cyclical Since The Mid-1980s?," *International Economic Review*, 49(1), 135–184.

Cesa-Bianchi, A. (2015): "VAR Toolbox," *sites.google.com/site/ambropo/*.

Champagne, J., and A. Kurmann (2013): "The Great Increase in Relative Volatility of Real Wages in the United States," *Journal of Monetary Economics*, 60(2), 166–183.

Chang, Y., and J. H. Hong (2006): "Do Technological Improvements in the Manufacturing Sector Raise or Lower Employment?," *American Economic Review*, 96(1), 352–368.

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

Cover, J. P., and S. K. Mallick (2012): "Identifying sources of macroeconomic and exchange rate fluctuations in the UK," *Journal of International Money and Finance*, 31(6), 1627–1648.

Dorn, D. (2009): "Essays on Inequality, Spatial Interaction, and the Demand for Skills," *Dissertation University of St. Gallen no. 3613*.

Foote, C. L., and R. W. Ryan (2014): "Labor-Market Polarization Over the Business Cycle," *NBER Macroeconomics Anual*, 29.

Francis, N., and V. A. Ramey (2005): "Is Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited," *Journal of Monetary Economics*, 52(8), 1379–1399.

——— (2009): "Measures of per Capita Hours and Their Implications for the Technology-Hours Debate," *Journal of Money, Credit and Banking*, 41(6), 1071–1097.

Gaggl, P., and S. Kaufmann (2015): "The Cyclical Component of Labor Market Polarization and Jobless Recoveries in the US," *Working Paper*.

Galí, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 89(1), 249–271.

Galí, J., and T. van Rens (2014): "The Vanishing Procyclicality of Labor Productivity and the Great Moderation," *Working Paper*.

Jaimovich, N., and H. E. Siu (2014): "The Trend is the Cycle: Job Polarization and Jobless Recoveries," *NBER Working Paper No. 18334*.

Jung, J., and J. Mercenier (2014): "Onmodeling task, skill and technology upgrading effects of globalization with heterogeneous labor," *Economic Modelling*, 39, 49–62.

King, M., S. Ruggles, J. T. Alexander, S. Flood, K. Genadek, M. B. Schroeder, B. Trampe, and R. Vick (2010): *Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]*. Minneapolis: University of Minnesota, 2010.

Lefter, A., and B. M. Sand (2011): "Job Polarization in the U.S.: A Reassessment of the Evidence from the 1980s and 1990s," *Economics Working Paper Series 1103, University of St. Gallen.*

Lucas, R. E. (1987): *Models of Business Cycles*. Blackwell.

Mallick, S. K., and R. M. Sousa (2016): "The Skill Premium Effect of Technological Change: New Evidence from the US Manufacturing Sector," *International Labour Review*, Forthcoming.

Meyer, P. B., and A. M. Osborne (2005): "Proposed Category System for 1960 – 2000 Census Occupations," *BLS Working Paper*, 383.

Mosbacher, R. A., and R. Ortner (1989): "The Relationship between the 1970 and 1980 Industry and Occupation Classification Systems," *Bureau of the Census, Technical Paper*, 59.

Sahin, A., J. Song, and B. Hobijn (2010): "The Unemployment Gender Gap During the 2007 Recession," *Current Issues in Economics and Finance*, 16.

Scopp, T. S. (2003): "The Relationship between the 1990 Census and Census 2000 Industry and Occupation Classification Systems," *Bureau of the Census, Technical Paper*, 65.

Shim, M., and H.-S. Yang (2015): "The Implication of Changes in Hours Fluctuations on Welfare Costs of Business Cycles," *Economics Letters*, 128, 75–78.

Smith, C. L. (2013): "The Dynamics of Labor Market Polarization," *Finance and Economics Discussion Series 2013-57. Board of Governors of the Federal Reserve System (U.S.).*

A Appendix

A.1 Aggregate Data Construction The March CPS and the MORG data report a person's primary occupation. Respondents who held more than one job were required to report the job in which they worked the largest number of hours. Following Autor (2010), we assign "Executive, administrative, and managerial occupations," "Professional specialty occupations," and "Technicians and related support occupations" to the high-skill occupation group; "Sales occupations," "Administrative support occupations, including clerical," "Precision production, craft, and repair occupations," "Machine operators, assemblers and inspectors," "Transportation and material moving occupations," and "Handlers, equipment cleaners, helpers and laborers" to the middle-skill occupation group; and "Private household occupations," "Protective service occupations," and "Service occupations except protective and household" to the low-skill occupation group. We exclude "Farming, forestry, and fishing occupations" and "Military occupations."[31]

The main variables of interest are constructed as follows:

1. *Employment*: In the CPS, individuals' employment status is determined on the basis of answers to a series of questions relating to their activities during the preceding week. Those who reported doing any work at all for pay or profit are deemed employed. We aggregate employment by their occupations in a given month or year with their sampling weight:

$$Employment_{D,t} = \sum_{i \in D} \mathbf{1}_{i,t}^{employed} u_{i,t} \tag{A.1}$$

   where $\mathbf{1}_{i,t}^{employed}$ indicates employment status, which equals one when individual i is employed at time t and zero otherwise. D is the individual's group category by occupation. $u_{i,t}$ is the individual sample weight.[32]

2. *Total hours worked*: We compute the total hours of work using the earnings weight provided by the CPS as follows.

[31] The categorization of workers in Cortes (2016) is almost identical to ours, but the author uses the terms "non-routine cognitive (high-skill occupation)," "routine (middle-skill occupation)," and "non-routine manual (low-skill occupation)."

[32] When aggregating individual data, we use the earnings weight (earnwt) that should be used in analyses of employment as well as hours/weeks worked.

$$TotalHours_{D,t} = \sum_{i \in D} h_{i,t} u_{i,t} \tag{A.2}$$

where $h_{i,t}$ is weekly hours worked for individual i at time t.

A.2 Additional Figures and Table from Other Data Sources We provide figures for employment and total hours worked from the March CPS in Figures A.1 to A.3. The findings in Section 3 can still be observed. For instance, there are two breaks for employment of the high-skill occupation group when the occ1990dd classification is used: between 1982 and 1983 and between 2002 and 2003. Similar breaks can be observed in Figure A.2b, total hours of the high-skill occupation group constructed using the occ1990dd classification. Hence, data constructed using conversion factors still outperform other data sets.

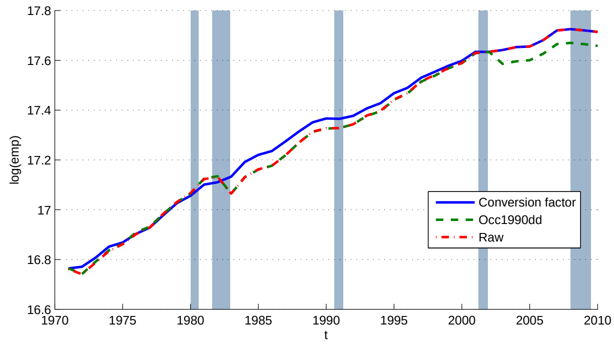

Figure A.1a: Employment, High-Skill Occupation

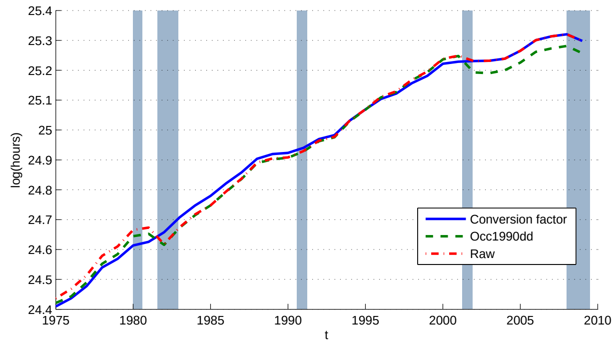

Figure A.1b: Total Hours, High-Skill Occupation

Figure A.1: March CPS

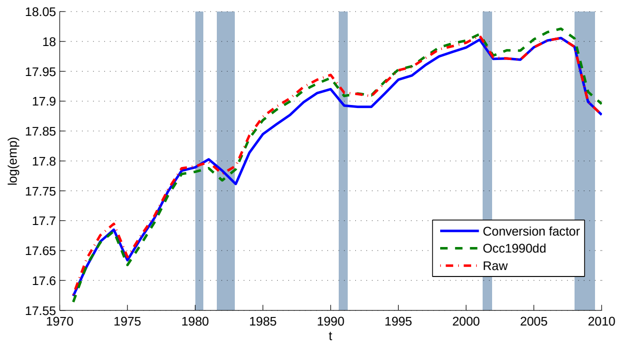

Figure A.2b: Employment, Middle-Skill Occupation

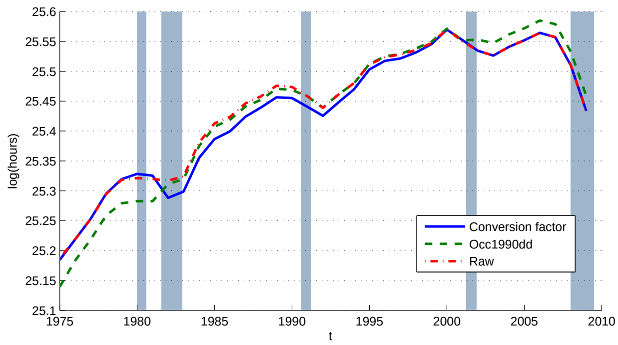

Figure A.2a: Total Hours, Middle-Skill Occupation

Figure A.2: March CPS

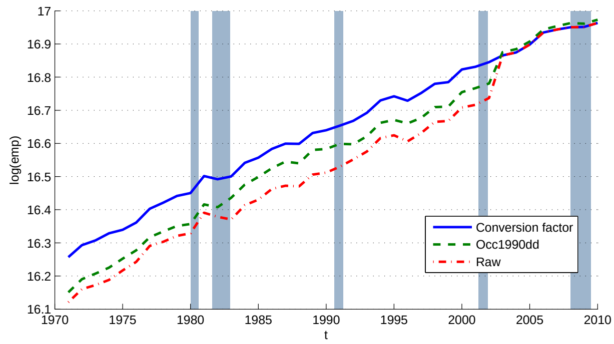

Figure A.3a: Employment, Low-Skill Occupation

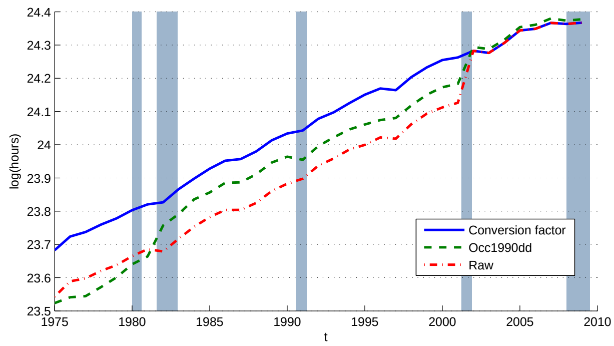

Figure A.3b: Total Hours, Low-Skill Occupation

Figure A.3: March CPS

Table A.1: Standard Deviations of Detrended Employment: March CPS

Occupation Group	High-Skill		Middle-Skill		Low-Skill	
Period	(1) 71–83	(2) 84–10	(1) 71–83	(2) 84–10	(1) 71–83	(2) 84–10
Standard Deviations	1.01	.93 (0.92)	2.34	1.10 (0.47)	1.13	0.84 (0.74)

Note: 1. All numbers are multiplied by 100.
2. Numbers in parentheses are the ratios between (2) and (1).

A.3 Changes in Employment Volatility: the Effect of Discontinuities in Data

To study the extent of the effect of discontinuities on standard deviations, we change the span of the first subperiods: 1979–1981, 1979–1982, and 1979–1983 (the benchmark case). The numbers in Table A.2 confirm our discussion. The volatilities of employment for the high-skill occupation group, which are constructed using the occ1990dd system or are from the raw data, decline dramatically as we shorten the first period to remove the effect of the discontinuities between 1982 and 1983. Note that our data set constructed using conversion factors still exhibits similar levels of standard deviations across the different periods, which comes from the fact that our data are consistent regardless of the changes in the coding scheme.[33]

A.4 Proof of Proposition 1

The model introduced here is a simplified version of Shim and Yang (2015). Superscript i to denote workers with different skills is omitted. There are two log-linearized equilibrium conditions:

[33] The increase in the standard deviation of employment for the middle-skill occupation group, which is a common feature across the three data sets, is the consequence of the 1981–82 recession, which mainly affected jobs that were included in the middle-skill groups.

Table A.2: Standard Deviations of Detrended Employment: Different Base Years

	79–81			79–82			79–83		
Methodology	H	M	L	H	M	L	H	M	L
Conversion Factors	0.29	0.71	0.33	0.30	0.75	0.44	0.39	0.84	0.51
occ1990dd System	**0.41**	0.69	0.33	**0.58**	0.92	0.33	**1.09**	0.88	0.47
Raw Data	**0.46**	0.76	0.40	**0.60**	0.90	0.52	**1.11**	0.87	0.56

Note: 1. All numbers are multiplied by 100.
2. H stands for high-skill, M stands for middle-skill, and L stands for the low-skill occupation group.

$$\hat{c}_t = \mathbb{E}_t \hat{c}_{t+1} \tag{A.3}$$

$$c\hat{c}_t + a\hat{a}_{t+1} = wh\hat{h}_t + (1+r)a\hat{a}_t \tag{A.4}$$

where $\hat{x}_t = \ln(x_t/x)$ with x being the steady state value. By assuming that $\hat{h}_t$ follows an AR(1) process, i.e. $\hat{h}_{t+1} = \rho_h \hat{h}_t + \varepsilon_{t+1}$ where $\varepsilon \sim \mathbb{N}(0, \sigma_h^2)$, we can guess $\hat{c}_t = \phi_0 \hat{h}_t + \phi_1 \hat{a}_t$ and $\hat{a}_{t+1} = \phi_2 \hat{h}_t + \phi_3 \hat{a}_t$, which can be verified as follows.

$$\phi_0 = \frac{rwh}{c(r+1-\rho_h)}, \ \phi_1 = \frac{ra}{c}, \ \phi_2 = \frac{wh}{\left(\frac{r}{1-\rho_h}+1\right)a}, \ \phi_3 = 1 \tag{A.5}$$

One can show by iteration that

$$\hat{c}_t = \phi_0 \hat{h}_t + \sum_{j=1}^{t} \Phi_j \hat{h}_{t-j} \tag{A.6}$$

where $\Phi_j = \phi_1 \phi_2 \phi_3^{j-1}$.

Using the assumption on $\hat{h}_t$, we can get the following expression:

$$\begin{aligned}\mathbb{V}(\hat{c}_t) &= \left[\phi_0 + \sum_{j=1}^{t}\left(\sum_{k=1}^{j} \Phi_k \rho_h^{j-k} + \phi_0 \rho_h^j\right)\right]\sigma_h^2 \\ &\propto \sigma_h^2 \end{aligned} \tag{A.7}$$

where the second line holds when the structure of the economy does not change.