

Is There A Tradeoff between Employer-Provided Health Insurance and Wages?

Liye Zhu*, Southern Methodist University
October 2005

Abstract

Though most of the literature in health insurance and the labor market assumes a tradeoff between employer-provided health insurance and wages, its empirical validity has not been established. Employing Current Population Survey 2004 data, this paper assesses the tradeoff hypothesis in a distributional analysis framework using stochastic dominance tests. In addition, it contributes to the previous literature by incorporating an indirect effect of health insurance on wages into the analysis. Health insurance not only directly affects wages, but also indirectly by improving individual productivity. The results confirm the existence of a tradeoff for full time working wives, and explain why the previous literature fails to do so.

JEL: C12, C14, D31, I11, J3

Keywords: Employer-Provided Health Insurance, Tradeoff, Distributional analysis, Stochastic Dominance Tests

* I am grateful to Professor Daniel Millimet for his profound supervision and continuous help. Thanks to Professor Nathan Balke, Professor Tom Fomby and Professor Yi Deng for helpful comments and advice. All errors are mine. Corresponding address: Liye Zhu, Department of Economics, Southern Methodist University, Dallas, TX 75275. Tel: 214-768-2339. Email: lzhu@mail.smu.edu.

I. Introduction

In contrast to most other developed countries, health insurance in the US is both provided and financed predominantly by employers, especially for working aged individuals. Current Population Survey 2004 data show that 63% of American adult population is covered by employer-provided health insurance (HI)¹ and this percentage changes only slightly over time. The magnitude of employer-provided health insurance coupled with the institutions and “rules” for health insurance provision have made health insurance an important parameter of labor market decisions both for individuals and firms. Theory predicts that there is a tradeoff between employer-provided health insurance and wages. Therefore, most studies of health insurance effects on both labor force participation and job choice as well as on other labor issues are based on this assumption. However, the link has not been empirically affirmed.

The focus of the recent empirical studies of the hypothesis tests has centered on problems with the data and the endogeneity of HI. However, even using "good" data and applying methods, such as panel data or instrumental variable (IV) models, to solve the endogeneity problem, researchers still cannot obtain strong results in support of the tradeoff hypothesis. Most find a positive relationship; a few obtain results that support the tradeoff hypothesis, but have certain shortcomings in their studies which weaken the results.

In addition to addressing these two difficulties, this paper also deals with two other possible shortcomings of the existing literature that may explain the inconsistent results. First, the basic linear regression model employed in most studies may be misspecified. Second, the effect of HI on wages may be heterogeneous.

In terms of model specification, this paper not only considers the direct effect of HI on wages, but also incorporates the indirect effect of HI into the analysis. HI may increase individual productivity, and therefore the returns to other determinants of productivity (such as education), by enhancing individual health or boosting working morale. Therefore, HI enters the wage equation, not only from the direct tradeoff point of view, but also from the productivity-enhancing point of view. The direct effect of HI can be reflected by the coefficient on HI in a wage equation. The indirect effect can be captured by the coefficients on the interactions between HI and

¹ Employer-provided health insurance is denoted as HI in this paper. Therefore, HI in this paper does not refer to whether individuals *have insurance or not*, but refers to whether individuals *have employer-provided health insurance or not*. Similarly, NOHI refers to individuals *without employer-provided health insurance*. It will be specified if health insurance status is in use.

individual or job characteristics in a wage equation. These two effects, the direct tradeoff effect and the indirect productivity-enhancing effect, work in opposite directions. The total effect of HI on wages depends on the magnitudes of these two effects. Since the coefficients on the interactions turn out to be significantly different from zero, and the interactions are highly correlated with other regressors, the previous literature which omits the interaction terms generates biased results.

Another possible reason that the previous literature cannot find significant tradeoff is the heterogeneity problem. If individual heterogeneity exists, the estimated tradeoff can be statistically insignificant when using regression analysis, which is based on the conditional mean, because the tradeoff will differ among individuals and may offset. To deal with individual heterogeneity, the tradeoff hypothesis is tested in this paper using a distributional analysis. The distributional analysis combined with tests of stochastic dominance (SD) differs from previous empirical regression analyses. It utilizes all available information to find and test for uniform rankings of distributions for different groups of people. Moreover, this approach allows one to easily assess potential heterogeneity in the magnitudes of tradeoff across the distribution of wages.

Specifically, the distributional analysis and stochastic dominance tests are performed in two cases: when HI is exogenous and when it is endogenous. In the first case, individuals are classified according to their HI status (HI or NOHI). Distributions of the following outcomes are obtained for each sample: the total wage, wages explained by the constant and residuals, and wages explained by individual and job characteristics. In addition, the hypothetical wages of NOHI sample using the HI regression coefficients are obtained. Then, the corresponding HI and NOHI distributions are compared by stochastic dominance tests. Wage differentials explained by the constant and residuals correspond to the direct effect in the typical regression analysis. By introducing the hypothetical wage, wage differentials explained by characteristics is further decomposed into two gaps: characteristic gap, which is solely due to differences in attributes, and compensation gap, which is due to different wage returns for the same characteristics across the two groups. The compensation gap corresponds to the indirect effect of HI, and is interpreted as improved individual productivity through better health or higher morale.

Since most of the previous literature predicts that HI is endogenous in the wage equation, it is worthwhile to deal with this endogeneity and thus the selection bias problem. When HI is endogenous, following the same

procedure as above, individuals are divided into two groups according to a binary instrument of HI. According to Abadie (2002), for the two groups divided by a binary IV, the difference between the cumulative distribution functions (CDF) is proportional to the difference between the CDFs of the compliers in each group under certain assumptions. Here, compliers refer to those individuals who comply with the IV, i.e., whose HI status changes according to the IV. Thus, we can test for stochastic dominance among the compliers by testing for dominance across the two groups divided by a binary IV. The instruments, borrowed from Olson (2002), are tested for their validity. The endogeneity of HI is also tested using the valid instruments.

The results in this paper, based on Current Population Survey (CPS) 2004 data on full-time working wives, not only suggest the existence of a tradeoff, but also suggest the existence of an indirect effect of HI on wages. This explains why previous studies fail to establish a tradeoff.

The paper is organized as follows: Part II is the literature review. Part III introduces the methods used in this paper. Part IV describes the data. Part V provides the results. Part VI concludes.

II. Literature Review

A. Theoretical Foundation of the Tradeoff Hypothesis

The theory of the tradeoff between wages and health insurance is that health insurance is a fringe benefit employers provide to employees as compensation. In a competitive product market, economic theory suggests that what matters to profit maximizing firms is the value of the total compensation package that they must offer to attract labor services. If the compensation level is too low, there will not be enough labor attracted to the firm; if the compensation level is too high, the firm cannot survive in the market. Thus, firms' compensation level to employees will be similar to that offered by other firms which face the same labor pool. As a result, if we assume that the other benefits offered by firms do not change, to remain competitive, firms will reduce wages by \$1 for each dollar increase in health insurance benefits. Individuals will then choose among firms offering different wage/health insurance combination according to their own preferences.

Under this scenario, there exists a tradeoff between wages and health insurance. As a benchmark, the predict tradeoff is $\partial W / \partial HI = -1$. The actual tradeoff would differ from this in certain situations. For example, the tradeoff should be less than -1 if the health insurance reduces job turnover and job turnover is a cost to the firm.

Another example, since health insurance costs are not taxable income for firms, but a cost before tax, the actual tradeoff should be $-1/(1-t)$, where t is the tax rate.

The figures in Currie and Madrian (1999) can illustrate the above theory. In Figure 1 (A), firm 1 and firm 2 are two firms facing a competitive product market and same labor pool. Employee 1 and Employee 2 need the same total compensation level, but have different preferences over wage compensation and health insurance compensation. We can see there is a tradeoff between wages and health insurance: if the wage level is high, health insurance will be low (Employee 2); on the other hand, a lower wage level is combined with a higher health insurance level (Employee 1). Thus, if all firms face the same tradeoff between wages and benefits in total compensation, the wage/health insurance bundles that are observed in the market will reflect the sorting of employees across firms on the basis of their heterogeneous preferences for health insurance. This framework is the motivation for much of the literature on the tradeoff between wages and employer-provided health insurance or other fringe benefits.

B. Empirical Tests of the Tradeoff Hypothesis

The empirical implementation of the wage-health insurance tradeoff pictured in Figure 1 (A) has typically been the estimation of

$$y_i = \alpha + X_i' \beta + HI_i \gamma + \mu_i \quad (1)$$

where y_i is labor market outcome of interest (here, log hourly wages) for individual i , X_i is a vector of individual and/or job characteristics for individual i , HI_i is either the availability or value of health insurance coverage for individual i , and μ_i is individual disturbance.

Conditional on X and in the absence of tax considerations, the theory predicts $\gamma = -1$ if y represents hourly wages and HI is appropriately measured in dollars. The empirical validity of Equation (1) with respect to wages, however, has been difficult to establish. The typical estimates of γ are either positive or insignificant. The literature has thus focused not on the magnitude of the wage-health insurance tradeoff, but on the reasons why economists cannot find evidence that there is one.

One main point that researchers focus on is the lack of a suitable dataset. In order to estimate Equation (1) data is required on both compensation and fringe benefit expenditures, that is, wage and health insurance levels. The firm-level datasets that include information on benefits expenditures are usually aggregated at the firm level, but typically do not include the types of human capital variables that might allow researchers to control for the productivity of the workforce. The problem created by these omitted variables is illustrated in Figure 1 (B). Employee 1 has higher abilities and thus earns more wages as well as fringe benefits, such as health insurance. Employer 2 has relatively low abilities and thus earns a lower wage and lower fringe benefits. Thus, if total compensation increases with average worker productivity and both benefits and other consumption goods are normal, a regression using such firm-level data will yield a positive relationship between wages and benefits. Some researchers thus merge in average employer expenditures by industry from a firm-level dataset to individual-level datasets. Even so, such methods still usually lead to a positive relationship. For example, Leibowitz (1983) uses the RAND Health Insurance Study to estimate the wage/fringe benefit tradeoff, but still finds a positive (although insignificant) effect of employer health insurance expenditures on wages.

Researchers then deduce that the reason that they cannot find a tradeoff is that productivity is determined by both observed human capital variables and unobserved (to the econometrician) ability. This can be shown as

$$y_i = \alpha + X_i' \beta + HI_i \gamma + \delta_i + \tilde{\mu}_i \quad (2)$$

where δ_i is some unobservable variable or fixed effect, $\tilde{\mu}_i$ is individual disturbance, and other variables are the same as in Equation (1). Thus, if δ_i is omitted and is correlated with health insurance, the estimation of γ will be biased. This implies that even conditional on observed human capital variables, some firms employ higher ability workers and pay a higher level of total compensation. But, as shown in Figure 1 (B), if this higher level of compensation is allocated to both wages and benefits, a positive relationship between wages and fringe benefits will be estimated despite using observable human capital controls.

Various approaches have been taken to circumvent this problem of omitted variable bias. One common method is the differencing method, i.e., purge the unobserved variable by differencing Equation (2), either across different years for same person, or across different job classifications within a firm. Buchmueller and Lettau (1997) use an employer-level dataset that tracks compensation and benefit expenditures for various jobs

within the firm over a 4-year period. They purge the unobserved productivity differences by differencing Equation (2) over time, essentially examining the impact of the growth in health insurance expenditures over time on changes in wages over time. Even so, they find no evidence of tradeoff between health insurance and wages. Olson (1992), Miller (1995) and Ryan (1997) use panel datasets of workers to estimate the effect of changes in health insurance coverage on changes in wages. However, the problem is that the majority of changes in health insurance coverage are generated by job changes and the unobserved job characteristics that also impact compensation are unlikely to be constant following a job change. The study by Olson is less subject to this criticism as his sample of displaced workers is exogenously selected by the closing of a plant or similar event. Gruber (1994) exploits a different source of variation, the changes of laws in many states in 1970s, which required employers who offered health insurance to treat pregnancy and childbirth the same as any other health issue. He finds that wages for those groups most likely to benefit from the law fell in direct proportion to the anticipated cost of the benefit. Overall, his results are consistent with a full shifting of employer health insurance costs onto wages.

A recent development in the literature to deal with the endogeneity problem is the IV method. Olson (2002), using Current Population Survey (CPS) 1990-1993 data, models the wages of married women employed full-time in the labor market. He uses husband's union status, husband's firm size, and husband's health insurance coverage through his job as instruments for wife's own employer-provided health insurance benefits. The estimates suggest that wives with own employer-provided health insurance accept a wage about 20% lower than what they would have received working in a job without benefits. However, Olson does not fully test the validity of the instruments, especially the orthogonality of the instruments with respect to wages.

III. Methodology

To test the tradeoff hypothesis, this paper not only considers the direct effect of HI on wages, but also incorporates the indirect effect of HI into the wage equation. In addition, this paper deals with individual heterogeneity problem and analyzes the tradeoff in a distributional framework. Finally, to solve the endogeneity due to omitted variables, the distributional analysis relies on the instrumental variables framework introduced by Abadie (2002).

A. Productivity Effect of HI

If we consider health as one kind of human capital, health insurance can be treated as an investment in this human capital. Employees typically start with a large health endowment that must be continuously replenished as it depreciates. If employees have employer-provided health insurance, they will have guaranteed health care, and their health risk can be alleviated. Therefore, HI can improve employees' health, which can then raise their productivity. Although employees with NOHI may still have other types of health insurance, HI is usually more generous and covers higher health care spending. Moreover, employees' productivity can also be enhanced by higher morale due to better health care. Thus, HI enters the wage equation not only from the direct tradeoff point of view, but also from the productivity-enhancing point of view.

Figure 1 (C) illustrates this productivity effect of HI. When employee 1 obtains better health insurance from her employer, her productivity improves and therefore the firm's profit increases. On the graph, this higher profit can be reflected by an iso-profit curve, which is higher than the original iso-cost curve for the firm. In a competitive market, the firm will pay higher wages to employee 1 to reflect the increase in profit. Therefore, the actual wage of employee 1 is not W_2 but W_3 . Movement from A to B reflects the direct tradeoff between HI and wages; movement from B to C reflects the indirect effect of HI on wages. These two effects work in opposite directions. The total effect of HI on wages is reflected by the movement from A to C, which only has an insignificant change of wages, since the above two effects offset each other.

In the regression analysis, this indirect effect of HI can be captured by adding the interaction terms between HI and other characteristic variables into equation (2)

$$y_i = \alpha + X_i'\beta + HI_i\gamma + HI_iX_i'\theta + \mu_i \quad (3)$$

where vector θ , the coefficients on the interaction terms, measures the different returns to the same characteristics for individuals with HI and with NOHI. As before, the coefficient on HI, γ , captures the HI direct effect. The previous regression analysis ignores the indirect effect and thus may result in biased estimation if $\theta \neq 0$, since the interaction terms are correlated with HI.

B. Distributional Analysis Assuming Exogeneity of HI

To access the degree of heterogeneity, this paper tests the tradeoff hypothesis using distributional analysis by comparing the wage distributions of wives with HI and wives with NOHI. This comparison between two distributions can be formalized using stochastic dominance (SD).

Several tests for SD have been proposed in the literature; the approach herein is based on a generalized Kolmogorov-Smirnov test. Suppose F_1 and F_0 are CDFs for two groups of individuals, group 1 and group 0, then the following definitions apply: F_1 First Order Dominates (FSD) F_0 if and only if $F_1(y) \leq F_0(y)$, $\forall y \in Y$, with strict inequality for some y ; F_1 Second Order Dominates (SSD) F_0 if and only if $\int_{-\infty}^y F_1(t)dt \leq \int_{-\infty}^y F_0(t)dt$, $\forall y \in Y$, with strict inequality for some y . The tests characterize the relationship between the distributions. Therefore, if the NOHI wage CDF dominates the HI wage CDF, there exists a tradeoff between employer-provided health insurance and wages.

To test for dominance, define the empirical CDFs for Y with HI and with NOHI respectively as

$$\hat{F}_{HI,N}(y) = \frac{1}{N} \sum_{i=1}^N I(Y_{HI,i} \leq y) \quad (4)$$

$$\hat{F}_{NOHI,M}(y) = \frac{1}{M} \sum_{j=1}^M I(Y_{NOHI,j} \leq y) \quad (5)$$

where N (M) is the size of the sample with HI (NOHI).

Now define the following functions of the joint distribution

$$d = \sqrt{\frac{MN}{M+N}} \min \sup_y [F_{NOHI,M}(y) - F_{HI,N}(y)] \quad (6)$$

$$s = \sqrt{\frac{MN}{M+N}} \min \sup_y \left\{ \int_{-\infty}^y [F_{NOHI,M}(t) - F_{HI,N}(t)] dt \right\} \quad (7)$$

where \min is taken over $F_{NOHI} - F_{HI}$ and $F_{HI} - F_{NOHI}$ (in effect performing two tests). The tests for FSD and SSD are based on the empirical counterparts of d and s (\hat{d} and \hat{s}) using the empirical CDFs. The test for FSD requires:

- (i) computing the values of $\hat{F}_{NOHI,M}(y_q)$ and $\hat{F}_{HI,N}(y_q)$ for $y_q, q = 1, \dots, Q$, where Q denotes the number of points in the support Y utilized ($Q = 500$ in the application),
- (ii) computing the differences $d_{1q} = \hat{F}_{HI,M}(y_q) - \hat{F}_{NOHI,N}(y_q)$, $d_{2q} = \hat{F}_{NOHI,M}(y_q) - \hat{F}_{HI,N}(y_q)$,
- (iii) finding $\hat{d} = \min\{\max\{d_{1q}\}, \max\{d_{2q}\}\}$.

If $\hat{d} \leq 0$ (to a degree of statistical certainty), then the null hypothesis of first order dominance is not rejected.

Furthermore, if $\hat{d} \leq 0$ and $\max\{d_1\} < 0$, then Y_{HI} FSD Y_{NOHI} as the value of the CDF for distribution Y_{NOHI} is at least as great as the corresponding value for distribution Y_{HI} at all $y_q, q = 1, \dots, Q$. On the other hand, if $\hat{d} \leq 0$ and $\max\{d_2\} < 0$, then Y_{NOHI} FSD Y_{HI} .

The analogous test for SSD requires:

- (i) computing the values of $\hat{F}_{NOHI,M}(y_q)$ and $\hat{F}_{HI,N}(y_q)$ for $y_q, q = 1, \dots, Q$, where Q denotes the number of points in the support Y utilized ($Q = 500$ in the application),
- (ii) computing the differences $d_{1q} = \hat{F}_{HI,M}(y_q) - \hat{F}_{NOHI,N}(y_q)$, $d_{2q} = \hat{F}_{NOHI,M}(y_q) - \hat{F}_{HI,N}(y_q)$,
- (iii) calculating the sums, finding $s_{1q} = \sum_{i=1}^q d_{1q,i}$, $s_{2q} = \sum_{i=1}^q d_{2q,i}$, $q = 1, \dots, Q$,
- (iv) finding $\hat{s} = \min\{\max\{s_{1q}\}, \max\{s_{2q}\}\}$.

If $\hat{s} \leq 0$ (to a degree of statistical certainty), then the null hypothesis of second order dominance is not rejected.

Moreover, if $\hat{s} \leq 0$ and $\max\{s_{1q}\} < 0$, then Y_{HI} SSD Y_{NOHI} as the cumulative value of the CDF for distribution Y_{NOHI} exceeds the corresponding value for distribution Y_{HI} at all $y_q, q = 1, \dots, Q$; otherwise, if $\max\{s_{2q}\} < 0$, then Y_{NOHI} SSD Y_{HI} .

This paper approximates the empirical distribution of the test statistics using bootstrap techniques. For each of 1000 bootstrap samples, \hat{d} and \hat{s} are computed. Thus, whether the empirical distributions are characterized by FSD or by SSD can be reported. The bootstrap reported probabilities represent the critical levels associated with the non-rejection region.

To this point, Y_{HI} and Y_{NOHI} have represented two unconditional variables. However, the magnitudes of wages are not only decided by whether the employee has HI or has NOHI, but also decided by other individual and job characteristics. Therefore, we need to separate the HI effect on wages from the effects of other variables to correctly measure the tradeoff. From Equation (3), the predicted wages respectively for HI and NOHI groups are

$$\hat{y}_{k,i} = \hat{\alpha}_k + X'_{k,i} \hat{\beta}_k + \hat{\varepsilon}_{k,i}, \quad k = HI, NOHI \quad (8)$$

The partial wage denotes the wage part explained by the estimated intercept and residuals for HI and NOHI,

$$y^d_{k,i} = \hat{\alpha}_k + \hat{\varepsilon}_{k,i}, \quad k = HI, NOHI \quad (9)$$

The CDFs of this partial wage for HI and NOHI groups, denoted as F^d_{NOHI} and F^d_{HI} , can be compared by dominance tests. The difference between the two distributions corresponds to the direct effect in the regression analysis, and can be interpreted as the tradeoff between HI and wages if HI is exogenous.

The part left in wage equation (8) stands for wages explained by characteristic variables, which is

$$y^{ch}_{k,i} = X'_{k,i} \hat{\beta}_k, \quad k = HI, NOHI \quad (10)$$

The corresponding CDFs, denoted as F^{ch}_{NOHI} and F^{ch}_{HI} , can also be compared by dominance tests. However, the difference between the two distributions reflects differences in X and differences in returns. We can decompose this part using a similar method as the standard Blinder-Oaxaca decomposition. For a regression, the standard Blinder-Oaxaca decomposition is in the form of the average value. Here, the author extends the standard Oaxaca decomposition to the distributional analysis. A hypothetical wage CDF for individuals with NOHI can be calculated using the estimated coefficients from the wage equation of individuals with HI, i.e., $F^h(X'_{NOHI} \hat{\beta}_{HI})$ (shorted as $F^h_{NOHI/HI}$). With this, decomposition similar to the standard Oaxaca-Blinder decomposition can be performed,

$$F^{ch}_{HI} - F^{ch}_{NOHI} = [F^{ch}_{HI} - F^h_{NOHI/HI}] + [F^h_{NOHI/HI} - F^{ch}_{NOHI}] \quad (11)$$

The first part of the expression on the right hand side is the difference in the earnings distributions explained by different characteristics of individuals with HI and NOHI. It can be called the characteristics gap. The second part is the difference due to the estimated parameters or wage structure. It is labeled the unexplained part of the

wage difference. In this paper, we call it the compensation gap, the wage return differences between HI and NOHI individuals with the same characteristics. It reflects the indirect effect of HI on wages.

In general, FSD and SSD are tested for 5 pairs of CDFs: the original data, the partial wages, the predicted wage explained by the characteristic variables, and the hypothetical wages. This paper also presents graphs of the horizontal difference between any two CDFs in consideration. Specifically, at every quantile, the wage difference between two CDFs can be calculated and can be plotted against the accumulated probabilities. By this, the pattern of the difference across the entire wage distributions can be analyzed.

C. Distributional Analysis When HI Is Endogenous

Since most of the previous literature predicts that HI is endogenous in the wage equation, and since the wage equation may not be correctly specified, it is worthwhile to test the endogeneity of HI. The reason is that if we divide the data according to an endogenous variable, for example, HI, regression and distributional analysis will suffer from selection bias. However, if we can find instruments for HI, the tradeoff between HI and wages can be put into the instrumental variables environment developed by Abadie (2002).

Let $Y_i(0)$ be the potential outcome for individual i without treatment, or with NOHI. Let $Y_i(1)$ be the potential outcome for the same individual with HI. Let Z_i be a binary instrument for HI_i . Denote $HI_i(0)$ the value that HI_i would have taken if $Z_i = 0$; $HI_i(1)$ has the same meaning for $Z_i = 1$. In practice, for any particular individual we can not observe both $HI_i(0)$ and $HI_i(1)$. Instead the realized treatment $HI_i = HI_i(1) * Z_i + HI_i(0) * (1 - Z_i)$ is observed. Similarly, only $Y_i = Y_i(1) * HI_i + Y_i(0) * (1 - HI_i)$ is observed.

Under the assumptions of: (1) independence of the instrument: $(Y_i(0), Y_i(1), HI_i(0), HI_i(1))$ is independent of Z_i ; (2) first stage: $0 < P(Z_i = 1) < 1$ and $P(HI_i(1) = 1) > P(HI_i(0) = 1)$; (3) monotonicity: $P(HI_i(1) \geq HI_i(0)) = 1$, Abadie has the following lemma for individuals whose treatment status is affected by variation in the instrument: $HI_i(0) = 0$ and $HI_i(1) = 1$ (the subpopulation of compliers):

$$F_1^C(y) - F_0^C(y) = \frac{E[I(Y_i \leq y) | Z_i = 1] - E[I(Y_i \leq y) | Z_i = 0]}{E[HI_i | Z_i = 1] - E[HI_i | Z_i = 0]} = K * (F_1 - F_0), \quad (12)$$

where $F_1^C = E[I(Y_i(1) \leq y) | HI_i(1) = 1, HI_i(0) = 0]$ and $F_0^C = E[I(Y_i(0) \leq y) | HI_i(1) = 1, HI_i(0) = 0]$, denotes the distributions of compliers; $F_1 = E[I(Y_i(1) \leq y) | Z_i = 1]$ and $F_0 = E[I(Y_i(0) \leq y) | Z_i = 0]$, the conditional distributions given $Z_i = 1$ and $Z_i = 0$; $K = 1/(E[HI_i | Z_i = 1] - E[HI_i | Z_i = 0]) < \infty$. Under assumption (2), we know that $K > 0$.

This lemma states that for a binary instrument, the difference between two distributions of compliers is proportional to the difference between the two distributions categorized by the instrument. So, if we use the instrument to solve the endogeneity problem and divide the individuals into two groups according to the binary instrument, comparing F_1 and F_0 provides the sign of the difference between F_1^C and F_0^C . This is different from the comparison of the original distribution of F_{HI} and F_{NOHI} . Moreover, since the IV results apply only to compliers, different instruments will generate different results.

The instruments used in this paper are borrowed from Olson (2002): husband's firm size, husband's health insurance, and husband's Union status. Olson (2002) qualitatively analyzes the economic validity of these three instruments, but does not statistically test their validity. For a variable to be the right instrument, it must be correlated with the endogenous variable, and orthogonal to the error process.

The relevance of the instrument variable can be tested by examining the first stage regression. The test statistics relate to the explanatory power of the instrument in the regression. A statistic commonly used is the R-squared of the first stage regression with the included exogenous variables partialled out. Alternatively, this may be expressed as the F-test of the joint significance of all the instruments in the first stage regression.

The orthogonality of the instruments can be tested using two stage least squares (2SLS) or general method of moment (GMM) techniques. If the disturbance is homoskedastic, the GMM estimator is equivalent to the 2SLS estimator. If it is heteoskedastic, the 2SLS estimator is inefficient but consistent, whereas the standard estimated covariance matrix is inconsistent. This paper tests the heteroskedasticity of the disturbance in wife's wage equation using White/Koenker nR2 statistic and Pagan-Hall general statistic. In the context of GMM, the orthogonality of the instrument may be tested via the commonly employed Hansen's J statistic. This statistic is

the value of the GMM objective function, evaluated at the efficient GMM estimator. The J statistic is distributed as χ^2 with degree of freedom equal to the number of overidentifying restrictions. A rejection of the null hypothesis implies that the instruments are not satisfying the orthogonal conditions required for their employment. This may be either because they are not truly exogenous, or because they are being incorrectly excluded from the regression.

Since the model is overidentified, testing a subset of the overidentifying restrictions is possible. In this context, the C test allows us to test a subset of the original set of orthogonality conditions. The statistic is computed as the difference between two J statistics: that for the (restricted, fully efficient) regression using the entire set of overidentifying restrictions, versus that for the (unrestricted, inefficient but consistent) regression using a smaller set of restrictions, in which a subset of instruments are removed from the set. The C statistic, distributed χ^2 with degrees of freedom equal to the loss of overidentifying restrictions, has the null hypothesis that the specified variables are proper instruments.

Using the valid instruments, this paper then tests whether wife's HI is exogenous in the wage equation. Both Durbin-Wu-Hausman (DWH) and the Wu-Hausman tests are applied. For the Durbin-Wu-Hausman test, a quadratic form in the differences between the two coefficient vectors, the IV estimator which is fully efficient under the null but inconsistent if the null is not true and the OLS estimator which is consistent under both the null and the alternative hypotheses, scaled by the precision matrix, gives rise to a test statistic for the null hypothesis that the OLS estimator is consistent and fully efficient. The test statistic is distributed as χ^2 with the degrees of freedom equal to the number of regressors being tested for endogeneity, which equals 1 in this paper. The asymptotically equivalent Wu-Hausman test is an F-test of the significance of the first stage residuals in the auxiliary second stage regression of 2SLS. One advantage of the Wu-Hausman F-statistic over the Durbin-Wu-Hausman test is that with certain normality assumptions, it is a finite sample test exactly distributed as F.

IV. Data

The data used in this paper come from the 2004 Current Population Survey (CPS) March Supplement dataset, and the 2004 CPS January Basic dataset.

The CPS is a monthly survey of a probability sample of housing units each month. The Annual Demographic Survey or March CPS Supplement is the primary source of detailed information on income and work experience in the United States. The labor force and work experience data from this survey are used to profile the U.S. labor market and to make employment projections. More importantly, the March CPS Supplement provides rich data on the individual health insurance information. CPS has a variable HI, which indicates whether or not the individual is covered by a health insurance plan provided through current/former employer/union.

However, the 2004 March Supplement does not include data on tenure: i.e., how long the worker has been in the current job; only the 2004 January Basic dataset has this information. So, this paper first merges these two data sets together. This step drops a lot of observations, due to the structure of the CPS.² Then, married individuals are chosen to match the husband and wife. From this couple dataset, the observations are further restricted to the following: individuals aged 25-60, employed full time, in the private sector, whose main earnings are wages and salary. The observations with the incomplete information³ are dropped, as well as individuals with hourly wages less than two dollars. After these restrictions, the sample size is 1287 couples.

The following are the brief introduction of the variables used in this paper.

The logarithm of current main job hourly wage is the dependent variable. This paper chooses the current main job earnings instead of total earning because currently a lot of part time individuals work several jobs, but the insurance coverage mainly comes from the main job. The CPS only provides the total wages and the earnings from the other jobs. So, subtracting the latter from the former yields the wages from the main job. The

² The CPS is a monthly survey of a probability sample of housing units each month. It does not, however, survey a completely new set of housing units each month. Rather, the sample is divided into eight representative sub samples called rotation groups, with housing units in each rotation group being interviewed for four consecutive months, followed by an 8-month break, and then by another four months of interviews. Thus, CPS sample housing units are each eligible for 8 different monthly interviews, and rotation groups are referred to in CPS parlance by their “month in sample” of MIS. In any given monthly sample, approximately one-eighth of sample units will be interviewed for the first time (MIS=1), one-eighth for the second time and so on. One eighth of the sample will leaving the sample permanently (MIS=8), and one-eighth of the sample will be leaving for the next eight months before being re-interviewed (MIS=4). These latter two rotation groups, MIS=8 and MIS=4, are referred to as the “outgoing rotation groups”. So, 75% of the CPS sample is common from month to month (any consecutive two months); while 50% of the CPS sample is common from one year to the next for the same month. However, because of one-response, mortality, migration and recording errors, there maybe still some errors after we match the two surveys using the household number and individual line number (which identifies the individual in the household). Madrian and Lefgren (1999) test the different matching method and give us better matching strategy. First match two datasets using household identifier (H_IDNUM), individual line number within the household (A_LINENO) and household number (H_HHNUM) which equals 1 in the initial interview and increased by 1 if the household is replaced by other in the next interview. After matching using H_IDNUM, A_LINENO and H_HHNUM, imposing additional merge criteria on gender, race and age. If gender and race differ in two surveys for the same individual, or if the difference of age in time t+1 and time t is less than -1 or greater than 3, we delete these observations.

³ Mainly includes those individuals without health insurance status information, without tenure information, and without union membership information.

CPS also provides the hours worked per week for the main job and weeks usually worked per year. Thus, we can obtain current main job hourly wage, and use the logarithm of it as the dependent variable.

HI status is the target variable, $HI = 1$ if individuals have HI and 0 if not. The variable HI means that the individual has employer-provided health insurance, and NOHI means not.

Race⁴, Age, Age square, Education, Geographical location, Number of kids under 18 and Husband's yearly earnings. All of these variables are categorical variables except age and husband's yearly earnings, which are continuous. Race has three categories: white, black, and other. Education is categorized into 7 groups: no diploma, high school diploma, some college, bachelor, master, professional schools including MD, and doctoral degree. The CPS classifies the states into four regions: Northeast, Midwest, South and West. The number of kids under 18 is classified into two categories: no kids under age 18 and some kids under age 18.

Major industry⁵, Major occupation, Tenure, Tenure square, Current job firm size, and Union membership. The CPS has 14 major industrial codes and our dataset covers 13. CPS has 11 major occupation codes, and ours includes 10. The excluded one is armed forces. Tenure information comes from the January Basic dataset, and is calculated as the number of years the individual have been employed in the current job. Firm size equals one if the number of the employees exceeds 100, zero otherwise. Lastly, union membership equals one if the individual is a union member or the current job is covered by a union, and zero otherwise.

Husband's HI, Husband's firm size), and husband's union membership are potential instruments for wife's HI status and are defined similarly as wife's variables.

Table 1 provides summary statistics. From the table we can see that wife's log hourly wage ranges from 0.80 to 5.93, the mean is 2.67. Furthermore, it shows that the percentage of wives having HI is relatively low (0.59) compared to this percentage of husbands (0.68).

Table 2 provides the correlation between wife's HI status and its instruments: husband's HI, husband's firm size and husband's union membership. This gives us a rough view of the relationships. From the table we have the following findings. First, all three instruments are negatively correlated with wife's HI. This confirms that these three variables are potential instruments. Second, husband's HI is strongly correlated with wife's HI;

⁴ Race turns out to be insignificant in the wage equation, so I run two sets of the regression: with and without race variables. Since the results are similar, I only provide the results without using this variable.

⁵ Industry also turns out to be insignificant in the wage equation, so I run two sets of the regression: with and without major industry code. Since the results are similar, I only provide the results without using this variable.

husband's firm size and union membership are less correlated. This is reasonable since the latter two variables are indirectly correlated with wife's HI: they are correlated with wife's HI through the correlation with husband's HI.

V. Results

A. Regression Results

To begin, tests for heteroskedasticity are provided, using the following regressors: HI, age, age squared, education, geographical location, number of kids under 18, major occupation, tenure, tenure squared and husband's yearly earnings.⁶ The White/Koenker nR2 test statistic is equal to 68.36 ($p = 0.00$). The Pagan-Hall general test statistic⁷ is equal to 72.15 ($p = 0.00$). Therefore, the tests reject the null hypothesis that the disturbance is homoskedastic. Thus, in the following, only the GMM results are presented.⁸

The instrumental relevance and orthogonality test results are illustrated in Table 3. The results are listed for different combinations of the three instruments: husband's firm size (F-size), husband's health insurance status (HHI), and husband's union membership (Union). In general, unlike Olson (2002), this paper finds that husband firm size and husband HI status are valid instruments, but husband's union membership is not.

Specifically, for the relevance test, the second column of Table 3 is the partial R-squared. The third column reports the F-statistic; the p-values are in brackets. We can see that the partial R-squared is between 0.01 and 0.10. The most relevant instrument is husband's health insurance, for which the partial R-squared is 0.07, followed by husband's firm size whose partial R-squared is 0.02 and husband's union status whose partial R-squared is 0.01. The F-statistics are large, and all the p-values equal 0.00. These results are consistent with our previous findings in Table 2.

For the orthogonality tests, Table 3 reports the Hansen's J statistics for GMM model. The p-values are in the brackets. We can see that the F-size and HHI combination has the lowest J, which is 0.60 ($p = 0.44$). Thus, we do not reject the hypothesis that F-size and HHI are orthogonal to the disturbance and that they are not in the

⁶ Industry and race are not statistically significant, so they are dropped from the regression.

⁷ The instruments used for this statistic are husband's firm size and husband's health insurance status, valid instruments known from the following test results.

⁸ The author also tests instrument validity for the 2SLS model, and finds that the results do not differ significantly.

regression equation. In contrast, models with instruments combined with Union have very high J statistics; the p-values are below 0.1. This suggests that the hypothesis that Union is orthogonal to the disturbance should be rejected.

Table 3 also reports the C-statistic⁹ of sub-instruments tests given the valid instruments. However, since the three instruments are highly correlated, we only have the C-statistics for F-size, HHI and Union given all the three instruments. We can see that F-size has the lowest C, followed by HHI and Union. For example, the C statistic for F-size is 0.41 ($p = 0.52$); for HHI the C statistics is 1.31 ($p = 0.25$); for Union it is 5.57 ($p = 0.02$). Therefore, the C-statistics also suggest that F-size and HHI pass the orthogonality test, but not Union.

Moreover, Table 3 lists the HI coefficients of the IV regressions. The coefficients are negative, which means that there are tradeoffs between employer-provided health insurance and wages. However, they are not statistically significant, except when using Union as instrument.

After testing the validity of the instruments, Table 4 reports the endogeneity test results using the valid instruments. The Wu-Hausman test statistics for F-size, HHI and both IVs are 2.19, 1.69 and 2.57 respectively; the p-values are 0.14, 0.19 and 0.11 respectively. The Durbin-Wu-Hausman tests generate similar results. Thus, we cannot strongly reject or accept the null hypothesis that the HI variable is exogenous, since the related p-values are all between 0.10 and 0.20, the margin of rejection and acceptance¹⁰. So, in the following, results assuming exogeneity and endogeneity are provided.

Finally, OLS and IV regression results are provided in Table 5, comparing the regressions with and without the interactions. For the regressions without interactions, column 2 of Table 5 is the HI coefficient of the OLS regression, with the p-value in brackets. Column 3 reports the HI coefficient of the GMM model, with the p-value in brackets. For the regression with interaction terms, column 4 and 6 reports the OLS and the GMM¹¹ coefficients respectively, with p-values in brackets; column 5 and 7 are the F-statistics of all the interaction terms. In the regressions with interactions, not only is HI instrumented, but also all the interactions, since they

⁹ The C statistic for a single instrument is 0, since there are no over-identifying restrictions for one instrument of one endogenous variable.

¹⁰ The author also test the Hausman test, plus Durbin's flavor of Hausman test and Wu's flavor of Hausman test. However, for these Hausman tests, the null hypothesis is that the differences of the coefficients of IV and OLS regression are not symmetric. So, strictly speaking, these are not endogeneity tests, but rather tests of the difference of two different regression methods: OLS vs. IV. So, although the tests suggest that the differences are not systematic (all the p-value are 1), we still cannot say that HI is exogenous.

¹¹ The author also estimates 2SLS models with the first stage linear, probit and logit respectively, but they do not generate fundamentally different results.

are also endogenous. Rows in Table 5 are for OLS model, the model with F-size as instrument, and the model with HHI as instrument respectively.

From the table, we have the following findings. First, there is large difference of the HI coefficients between whether or not we include the indirect effect of HI. For example, for OLS regressions, the coefficient on HI excluding the interactions is 0.09, while with the interactions, it is -0.06. Though it is not statistically significant, the sign has changed from positive to negative. For the IV regression, the HI coefficient is much bigger in absolute value for the model with interactions than the model without. For example, HI coefficients in the IV regression with the F-size as the instrument are -0.22 and -1.82 respectively for the non-interaction and interaction regression. They are -0.06 and -1.09 for HHI as the instrument.

Second, for the regressions including the indirect effect of HI, the coefficients on the interactions are statistically different from zero for some models. For example, in the IV model with HHI as the instrument, the coefficients on interaction terms are statistically different from zero. Therefore, HI affects wages not only from the direct tradeoff point of view, but also from changing the returns of the individual characteristics.

Third, although we have the expected sign of the coefficient on HI, they are not statistically significant. Only the HI coefficient of the OLS regression without interaction terms is statistically significant at 10% level. All the other coefficients are not statistically significant. These results lead us to do the distributional analysis.

B. Distributional Analysis Results

This part reports the distributional analysis results when HI is exogenous and when it is endogenous. Figures 2-4 provide the CDF comparisons in graph. Figure 2 is the CDF comparisons assuming that HI is exogenous in wife's wage equation, and therefore the individuals are divided into HI and NOHI groups by the HI variable. Figure 3 and 4 are CDF comparisons when wife's HI is endogenous in her wage equation. Here, individuals are divided into HI and NOHI groups by binary instruments. Figure 3 is the results when using husband's firm size as the instrument. Figure 4 is the results when using husband health insurance as the instrument. For each figure, graphs in the first column are CDFs; graphs in the second column are wage differentials of CDFs at different quantiles; graphs in the third column are accumulated CDFs. The graphs in the first row are for total wages; graphs in the second row are for the partial wages, which reflect the direct effect of HI; and graphs in the third

row are for the predicted wages explained by individual characteristics ($X\hat{\beta}$), as well as the hypothetical wages for individuals with NOHI using the estimated coefficients for individuals with HI.

Corresponding to these graphs, Table 6 also provides the quantitative levels of these CDFs and differentials at the mean and at quantiles 10, 25, 50, 75 and 90. The three rows are results for the following different models respectively: without instrument, husband's firm size as the instrument and husband's health insurance status as the instrument. The three columns are results of total wage and wage differentials, wages explained by the constant and residuals as well as wage differentials, and the wages explained by observable characteristics ($X\hat{\beta}$) and Oaxaca decomposition results.

Table 7 provides the stochastic dominance test results corresponding to the graphs in Figures 2-4. The three rows are results for the following different models respectively: without instruments, husband's firm size as instrument and husband's health insurance status as instrument. For each model, it also provides in detail the stochastic dominance test results for the total wage, for the partial wage, and the wage explained by the characteristic variables, which is sub-categorized as $X\hat{\beta}$ gap, X gap and the $\hat{\beta}$ gap. The $X\hat{\beta}$ gap stands for the total gap explained by the characteristic variables; the X gap is the wage gap solely due to the characteristic differences; $\hat{\beta}$ gap is the gap due to the compensation difference, the indirect effect of HI. FSD and SSD test results are based on 1000 bootstrap repetitions. d_{1MAX} , d_{2MAX} , d , $\Pr(d_1^* \leq 0)$, $\Pr(d_2^* \leq 0)$, $\Pr(d^* \leq 0)$ statistics are provided for FSD test; s_{1MAX} , s_{2MAX} , S , $\Pr(s_1^* \leq 0)$, $\Pr(s_2^* \leq 0)$, $\Pr(s^* \leq 0)$ statistics are provided for SSD test. The subscript 1 stands for the HI minus NOHI; the subscript 2 stands for NOHI minus HI. The superscript * stands for the results for each bootstrap. Furthermore, the larger $\Pr(d^* \leq 0)$ is, the more possible that there is FSD; and the larger $\Pr(s^* \leq 0)$ is, the more possible that there is SSD.

The following are findings from the graphs in Figure 2-4 and Tables 6-7.

B. 1. Total Wages

From the graphs of the first row in Figures 2-4 we can see that the CDFs of total wages do not reflect the tradeoff if we do not use an instrument. In fact, the CDFs of HI even dominate the CDFs of NOHI. Using the IV

approach, the situation reverses. The first column in Table 6 clearly implies this. The total mean wage difference is 0.26 when not using instrument, but ends up with -0.12 and -0.07, respectively, when using husband's firm size and health insurance as instruments. Table 7 lists the stochastic dominance results. Except NOHI second order dominates HI when using husband's firm size as instrument, the other relationships using instruments are not statistically significant. This finding explains why the previous researchers cannot find the tradeoffs when they do not use proper instruments.

B. 2. Partial Wages

From the second row graphs of partial wages in Figures 2-4, we can see that there exists a tradeoff between employer-provided health insurance and wages, which is the direct effect of HI on wages. When assuming the exogeneity of HI, the tradeoff is very small. From Table 6 we can see that this mean wage difference explained by constant and residuals is 0.02 in absolute value. When using husband's firm size as instrument, the difference is 0.24 in absolute value. When using husband's health insurance status as instrument, the difference is 0.51 in absolute value. However, these numbers depend on the magnitude of K , so they do not reflect the true tradeoff. Nevertheless, their signs indicate that there is a tradeoff between employer-provided health insurance and wages for compliers.

The stochastic dominance test confirms the above direct findings from the graph. Table 7 suggests that for the partial wages, there is no first or second order stochastic dominance for HI and NOHI groups when not using any instrument. For example, $\Pr(d^* \leq 0)$ is 0.024 and $\Pr(s^* \leq 0)$ is 0.454. When using the instruments, the NOHI not only second order stochastic dominates HI, but also first order dominates. For example, $\Pr(d^* \leq 0)$ is 0.932 and $\Pr(s^* \leq 0)$ is 0.993, and d_{2MAX} is -0.020 when using husband's firm size as instrument. When using husband's health insurance status as instrument, these three statistics are 0.999, 1.000 and -0.029 respectively.

Table 6 also suggests that the total wages are explained half by the characteristic variables and half by the constant and residuals. This is consistent with the literature, since the R-squared for wage equations are at most 0.5.

B. 3. Characteristic / Compensation Gaps

For the wage explained by the characteristic variables, the compensation gap which reflects the indirect effect of HI is positive and much larger than the characteristic gap. From the third row graphs in Figures 2-4, we can see that the hypothetical CDFs of NOHI group using HI regression coefficients are very close to the CDFs of the HI group, and far away from the CDFs of NOHI group. The last three columns of Table 6 illustrate this quantitatively. Except the one without instrument, the numbers in $\Delta W_{nohi/hi-nohi}$ column (compensation gap) are positive and almost equal to the numbers in $\Delta W_{hi-nohi}$ column (total gap); sometimes they are even larger. However, the numbers in the $\Delta W_{hi-nohi/hi}$ column (characteristic gap) are much smaller. This finding tells us that the characteristic differences for HI and NOHI individuals are not that big, at least for the characteristics we can observe. For example, when using husband's firm size as instrument, the mean wage difference due to the compensation difference is 0.20, while the total gap is 0.09 and gap due to the characteristic difference is -0.11. While for husband's health insurance status as instrument, the mean wage difference due to compensation difference is 0.50 and the total gap is 0.47 and the gap due to the characteristic difference is -0.03.

The test results in Table 7 confirm this finding. We can see from Table 7 that there are FSDs and SSDs for the $\hat{\beta}$ gap, the compensation gap or the indirect effect of HI, but no obvious FSDs for the X gap, i.e. the characteristic gap, when using instruments. For example, when using husband's firm size as instrument, $\Pr(d^* \leq 0)$ for $\hat{\beta}$ gap is 0.995, $\Pr(s^* \leq 0)$ is 0.996 and d_{1MAX} is equal to -0.024. However, these three statistics for X gap are respectively 0.384, 0.981 and -0.001. With husband's health insurance status as instrument, the three statistics for $\hat{\beta}$ gap are 0.999, 1.000 and -0.028 respectively. They are respectively 0.044, 0.710 and 0.017 for X gap.

This finding means that the wage returns are higher for individuals with employer-provided health insurance than those individuals without. In other words, the characteristics are not too different for HI and NOHI groups, at least for the characteristics we can observe, but the wage return of the same characteristics is very different for the two groups. The individuals with employer-proved health insurance not only get their health insurance from their companies, they also get higher pay. One possible explanation is that HI improves individual's productivity through investment in health; another explanation is that employees' morale is higher

when their employer providing health insurance. Therefore, HI affects wages not only from the direct tradeoff point of view, but also from the indirect productivity-enhancing point of view. These two effects partially offset, and the general effect depends on the magnitude of both effects.

B. 4. Other Findings

We should have noticed that the results are different when using different instruments. Since different sub-populations may respond differently to different instruments, the results will differ thereafter. Fortunately, the difference does not fundamentally alter the results.

Another way to analyze the results is along the distribution. When using different instruments, the wage differentials of two CDFs have different trends along the distribution. For example, considering the direct tradeoff effects, we can see from the graphs in second column second row of Figures 2-5, the differentials of predicted wages explained by constant and residuals have different trends. The differentials are almost constant when no instrument is used, except that they tilt down at the higher quantiles; when using husband's firm size as instrument, the situation reverses. They keep constant, except that they tilt down at the lower quantiles; when using husband's health insurance, the wage differentials have clear downward sloping trend.

The above trend differences tell us that the compliers not only behave differently in a mean perspective when using different instruments, but also behave differently at different wage levels. So, the regression analysis even using valid instruments may mask the tradeoff results. These findings shed some light on the policy making. For instance, when policy makers consider any policy related to the employer-provided health insurance and wages, they should consider the different reactions from different groups of people.

VI. Conclusion

Most of the studies of health insurance's effects on both labor force participation and job choice as well as other labor issues assume a tradeoff between employer-provided health insurance and wages, although the hypothesis does not have much empirical support from the previous research.

This paper incorporates the indirect effect of HI on wages: HI increases productivity and thus wage returns through better health and higher morale. Moreover, it utilizes the distributional analysis method to solve the

heterogeneity problem, and employs the instrument variable models to deal with the endogeneity problem. The author not only finds that there is a tradeoff between employer-provided health insurance and wages, but also suggests the existence of the indirect effect.

The results explain why previous researchers cannot establish the tradeoff. HI not only affects wages from the direct tradeoff point of view, but also affects wages by increasing the productivity through better health and higher morale of employees. These two effects work in the opposite direction and are closely related. Therefore, ignoring the indirect effect will result in biased estimation of the tradeoff.

The findings of this paper shed some light on policymaking. When policymakers consider a policy mandating a fringe benefit for workers, if the increase in fringe benefits entails a dollar-for-dollar tradeoff with wages, then the policy will do little to improve workers' utility (unless the policy counteracts some market failure), and may actually make workers worse off. However, given the presence of the indirect effects of expanding fringe benefits, such a policy is much more likely to be beneficial to most workers.

References:

- [1] Abadie, Alberto (2002), "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97, 284-292.
- [2] Baum, Christopher F., Mark E. Schaffer and Steven Stillman (2003), "Instrumental Variables and GMM: Estimation and Testing," *The Stata Journal*, 3, 1-31.
- [3] Buchmueller, Thomas C., John DiNardo and Robert G. Valletta (2002), "Union Effects on Health Insurance Provision And Coverage in The United States," *Industrial and Labor Relations Review*, 55, 610-627.
- [4] Buchmueller, Thomas C. and Michael K. Lettau (1997), "Estimating the Wage-health Insurance Tradeoff: More Data Problems?" *Working paper*, University of California at Irvine.
- [5] Currie, Janet and Brigitte C. Madrian (1998), "Health, Health Insurance and the Labor Market," *JCPR Working Paper 27*, Northwestern University/University of Chicago Joint Center for Poverty Research.

- [6] Fichtenbaum, Rudy and Paulette Olson (2002), "The Impact of Unionization on Health Insurance Benefits," *Journal of Economic Issues*, 36, 323-330.
- [7] Gruber, Jonathan (1994), "The Incidence of Mandated Maternity Benefits," *The American Economic Review*, 84, 622-641.
- [8] Gruber, Jonathan and Maria Hanratty (1995), "The Labor Market Effects of Introducing National Health Insurance: Evidence From Canada," *Journal of Business and Economics Statistics*, 13, 163-174.
- [9] Leibowitz, Arleen (1983), "Fringe Benefits in Employee Compensation," In *The Measurement of Labor Cost*, ed. Jack E. Triplett. Studies in Income and Wealth, No. 48. Chicago: University of Chicago Press for the National Bureau of Economic Research.
- [10] Linton, Oliver, Esfandiar Maasoumi and Yoon-Jae Wang (2005), "Consistent Testing for Stochastic Dominance under General Sampling Schemes," *Review of Economic Studies*, 72, 735-765.
- [11] Madrian, Brigitte C. and Lars John Lefgren (1999), "A Note on Longitudinally Matching Current Population (CPS) Respondents," *Technical Working Papers* 0247, NBER.
- [12] Miller, Richard D. (1995), "Estimating Compensating Differentials For Employer-provided Health Insurance Benefits," *Working paper*, University of California at Santa Barbara.
- [13] Olson, Craig A. (1992), "The Impact of Permanent Job Loss on Health Insurance Benefits," *Working paper*.
- [14] Olson, Craig A. (1998), "A Comparison of Parametric and Semiparametric Estimates of the Effect of Spousal Health Insurance Coverage on Weekly Hours Worked By Wives," *Journal of Applied Econometrics*, 13, 543-565.
- [15] Olson, Craig A. (2002), "Do Workers Accept Lower Wages In Exchange for Health Benefits?" *Journal of Labor Economics*, 20, 91-114.
- [16] Ryan, S. (1997), "Employer-provided Health Insurance and Compensating Wage Differentials: Evidence from The Survey of Income and Program Participation," *Working paper*, University of Missouri-Columbia.

Figure 1: Tradeoffs between employer-provided health insurance and wages

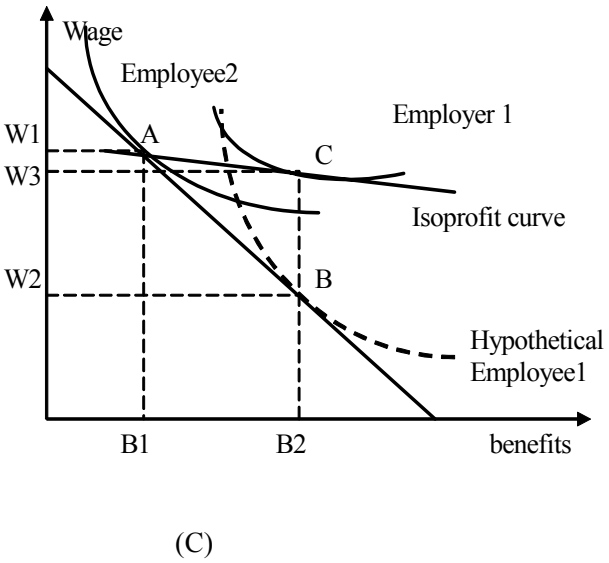
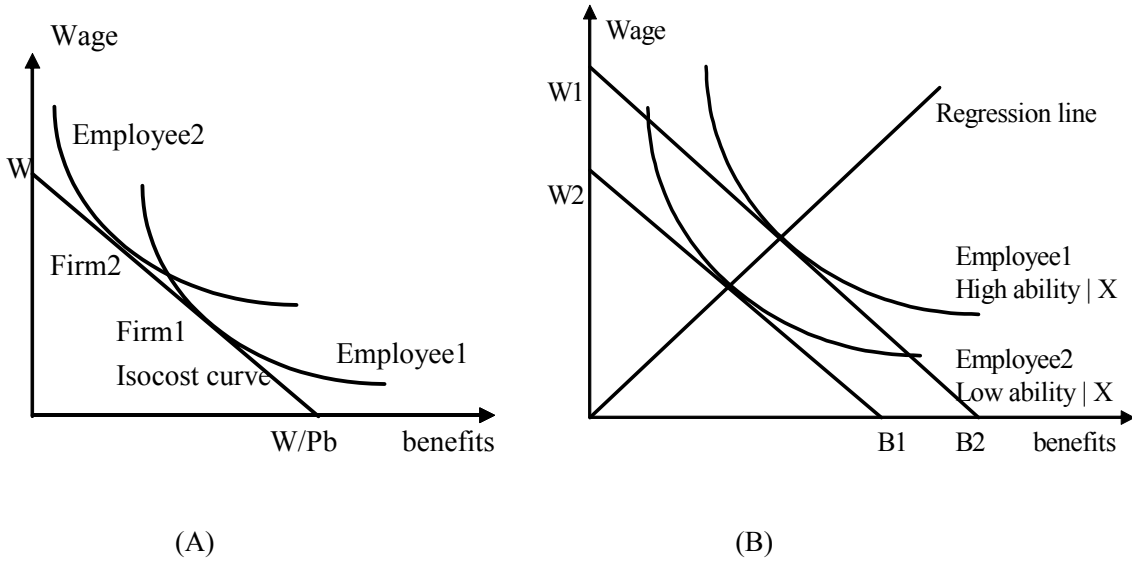
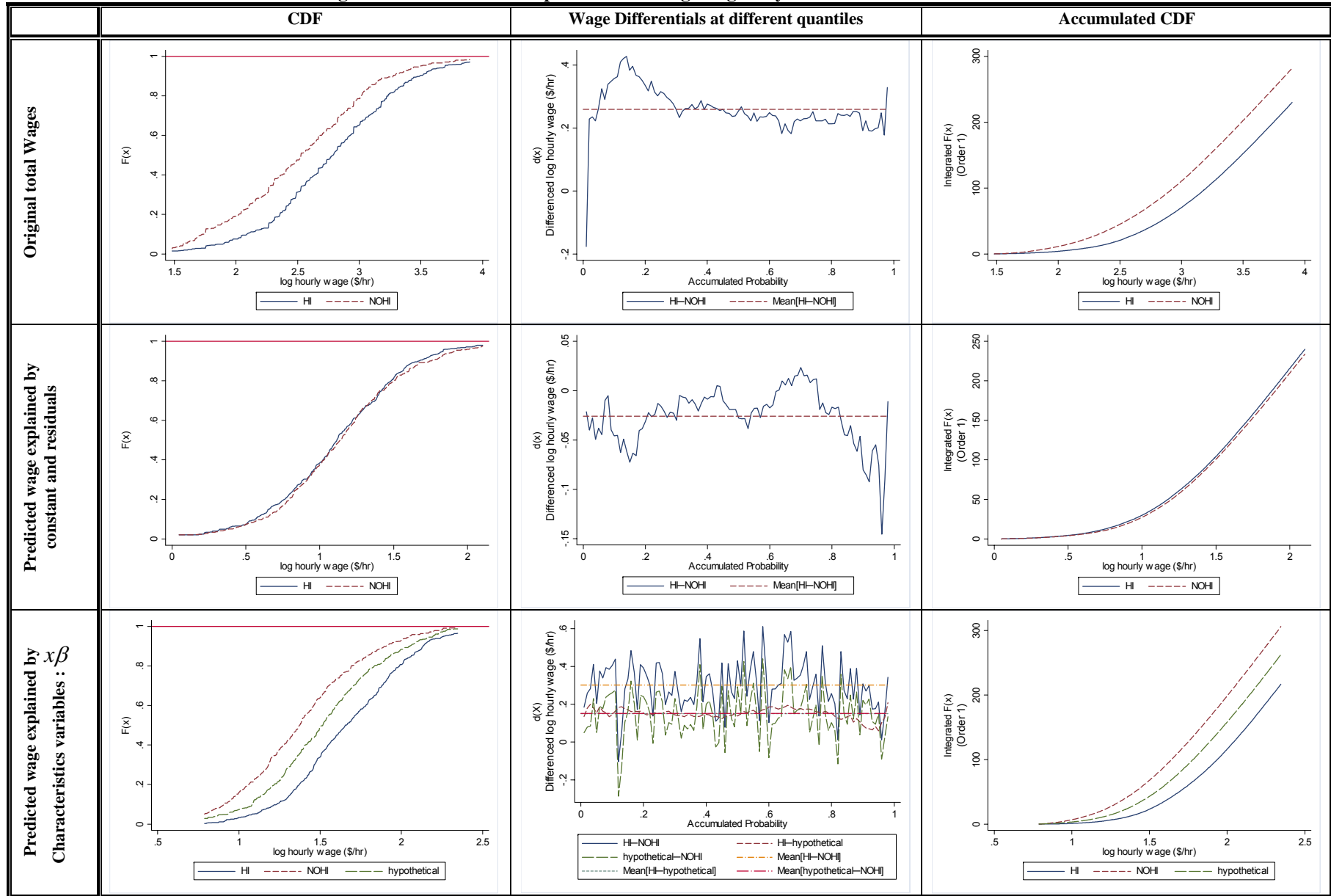
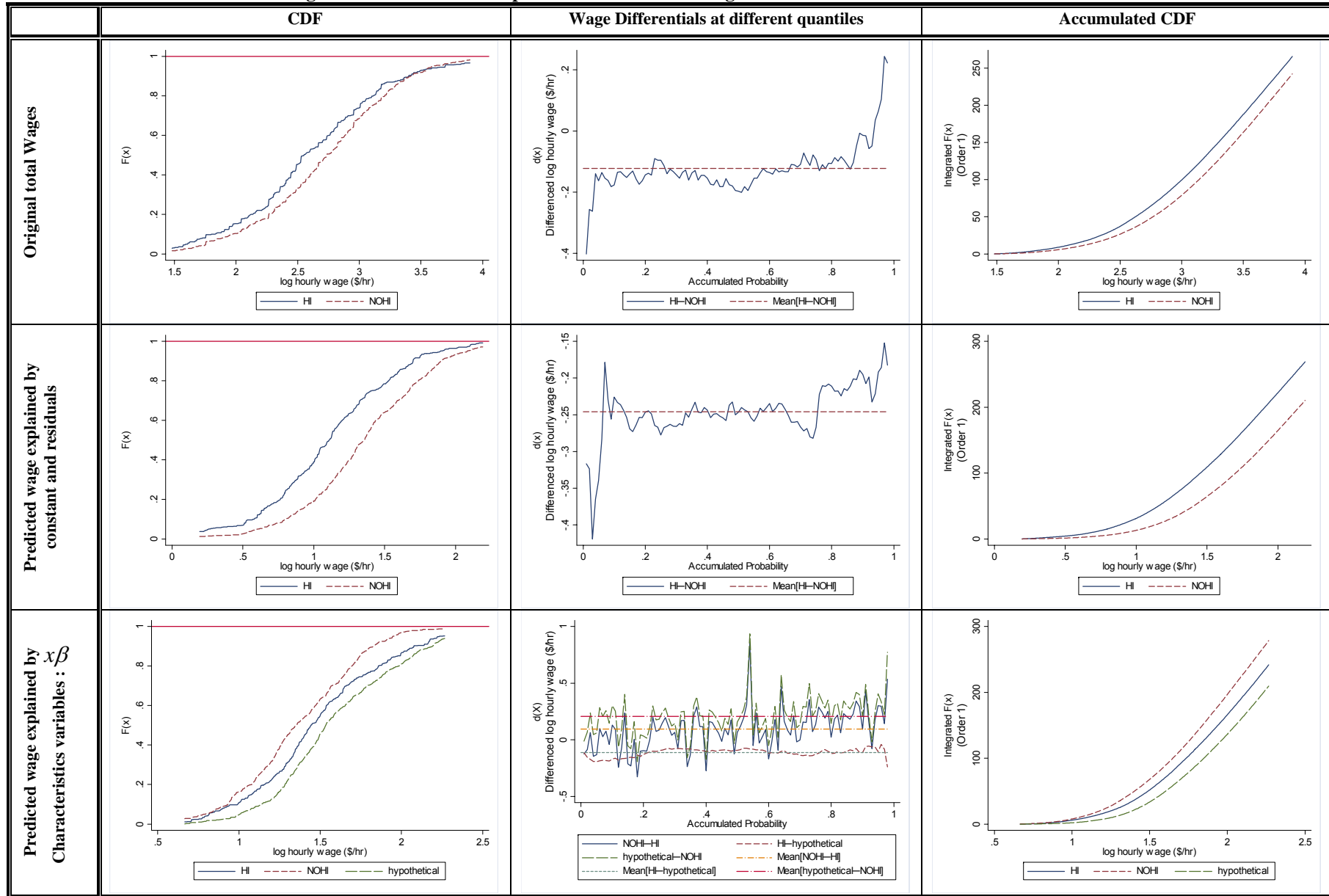


Figure 2: Distribution Comparison Assuming Exogeneity of HI - Without Instruments



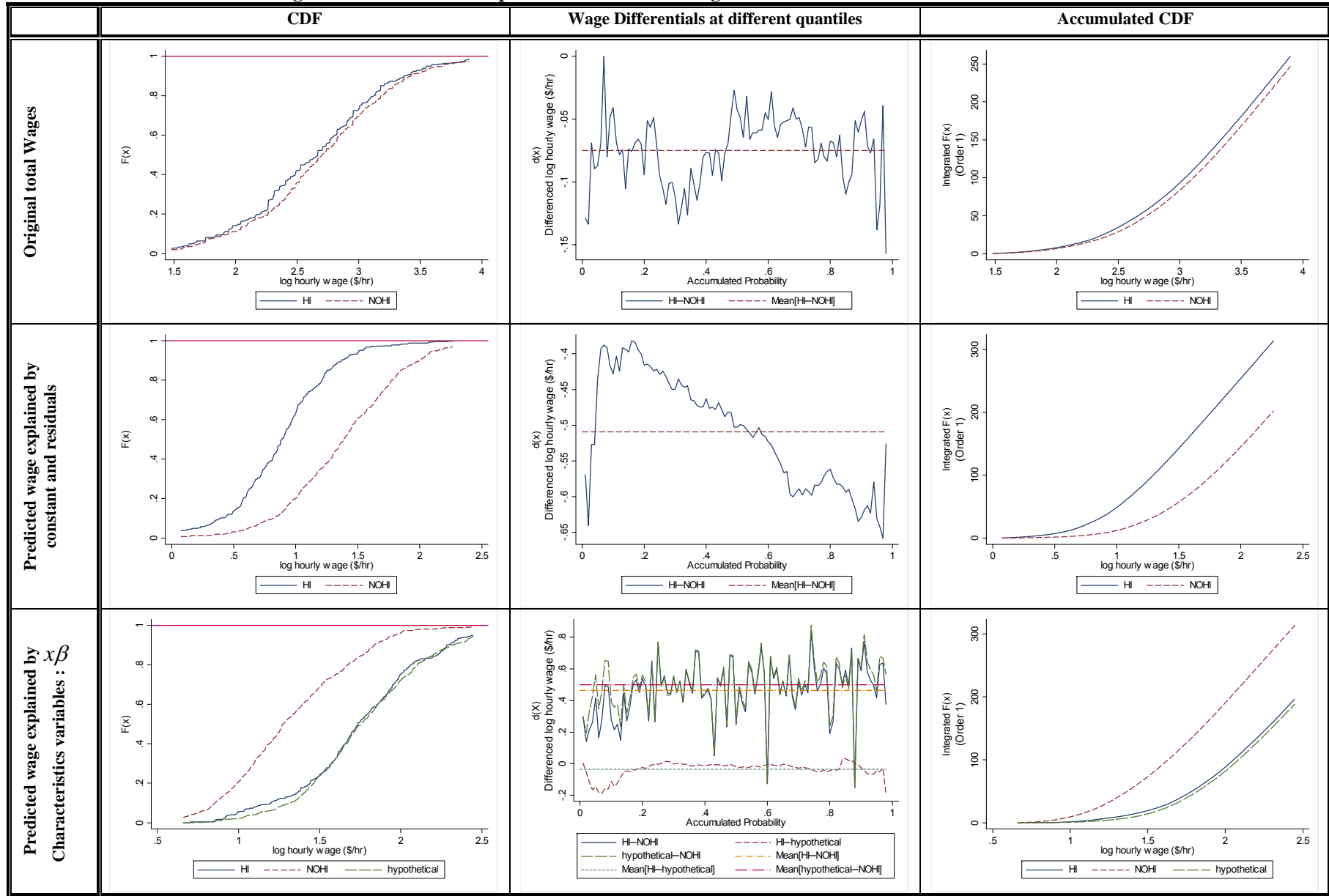
Note: The hypothetical CDF is the CDF of the predicted log hourly wages for NOHI group using the estimated coefficients of HI group.

Figure 3: Distribution Comparison When HI IS Endogenous - Husband's Firm Size as Instrument



Note: The hypothetical CDF is the CDF of the predicted log hourly wages for NOHI group using the estimated coefficients of HI group.

Figure 4: Distribution Comparison When HI IS Endogenous - Husband's Health Insurance as IV



Note: The hypothetical CDF is the CDF of the predicted log hourly wages for NOHI group using the estimated coefficients of HI group.

Table 1: Summary Statistics (total sample: 1287 couples)

Variable Name	Mean	Std. Dev.	Variable Name	Mean	Std. Dev.
Dependent Variable			Wife's Major Industry		
Wife's Log Hourly Wages	2.67	0.60	Agriculture / forestry / fishing / hunting	0.002	0.04
Independent Variables			Mining	0.002	0.04
Wife's HI Status			Construction	0.02	0.13
NOHI	0.41	0.49	Manufacturing	0.13	0.34
HI	0.59	0.49	Wholesale and retail trade	0.16	0.37
Wife's Race			Transportation and utilities	0.03	0.16
White	0.85	0.35	Information	0.02	0.16
Black	0.07	0.26	Financial activities	0.16	0.36
Other	0.08	0.27	Professional and business services	0.10	0.30
Wife's Age	40.35	8.48	Educational and health services	0.29	0.46
Wife's Geographical Location			Leisure and hospitality	0.05	0.22
Northeast	0.22	0.41	Other services	0.02	0.15
Midwest	0.31	0.46	Public administration	0.006	0.08
South	0.27	0.44	Wife's Major Occupation		
West	0.21	0.41	Management / business / financial	0.19	0.40
Wife's Education			Professional and related occupations	0.22	0.42
No diploma	0.04	0.21	Service	0.11	0.32
High school diploma	0.33	0.47	Sales and related occupations	0.11	0.31
Some college	0.33	0.47	Office and administrative support	0.27	0.44
Bachelor's degree	0.23	0.42	Farming / fishing / forestry	0.002	0.04
Master's degree	0.05	0.23	Construction/extraction	0.00	0.00
Professional school degree	0.01	0.08	Installation / maintenance / repair	0.006	0.08
Doctorate degree	0.01	0.09	Production	0.06	0.25
Wife's Tenure (year in the current job)	6.85	6.72	Transportation and material moving	0.02	0.15
Wife's Current Job Firm Size			Instrument Variables		
Small firm (# employees<100)	0.36	0.48	Husband's HI Status		
Big firm (# employees≥100)	0.64	0.48	NOHI	0.32	0.47
Wife's Union Membership			HI	0.68	0.47
Non union member	0.94	0.24	Husband's Firm Size		
Union member	0.06	0.24	Small firm (# employees<100)	0.38	0.48
Number of Kids Under 18			Big firm (# employees≥100)	0.62	0.48
Less than 1	0.37	0.48	Husband's Union Membership		
Greater or equal to 1	0.63	0.48	Non union member	0.87	0.33
Husband's Yearly Earnings	49789.3	44671.41	Union member	0.13	0.33

Table 2: Correlation between wife's HI status and its instruments (total sample: 1287 couples)

	Wife's HI	Husband's HI	Husband's F-size	Husband's Union
Wife's HI	1.00			
Husband's HI	-0.26	1.00		
Husband's F-size	-0.07	0.24	1.00	
Husband's Union	-0.09	0.14	0.09	1.00

Note: HI: employer-provided health insurance; F-size: firm size; Union: union membership.

Table 3: Instrument Validity Tests

Instruments tested	Relevance Tests		Orthogonality Tests				
	Partial R2	F Stat.	HI-Coef.	Hansen's J Stat.	C Statistic		
					F-size	HHI	Union
F-size, HHI, Union	0.08	33.90 (0.00)	-0.01 (0.94)	6.18 (0.05)	1.31 (0.25)	0.41 (0.52)	5.57 (0.02)
F-size, HHI,	0.07	47.64 (0.00)	-0.08 (0.49)	0.60 (0.44)	n/a	n/a	
F-size, Union	0.03	18.21 (0.00)	0.08 (0.63)	5.83 (0.02)	n/a		n/a
HHI, Union	0.07	47.13 (0.00)	0.02 (0.88)	4.90 (0.03)		n/a	n/a
F-size	0.02	24.23 (0.00)	-0.22 (0.32)	0.00			
HHI	0.07	86.95 (0.00)	-.057 (0.62)	0.00			
Union	0.01	15.27 (0.00)	0.53 (0.07)	0.00			

Notes:

1. F-size: husband's firm size; HHI: husband's health insurance status; Union: husband's union membership.
2. The p-values are in brackets. For Hansen's J statistic, the larger the p value, the more possible to accept the joint null hypothesis that the excluded instruments are valid instruments, i.e., uncorrelated with the disturbance, and that they are correctly excluded from the estimated equation.
3. The C statistic for a single instrument is 0, since there are no over-identifying restrictions if there is one instrument for one endogenous variable.
4. Due to highly correlation between any two of the instruments, the C statistic using any two instruments fails.
5. The regressors except HI in the wage equation are age, age square, education, geographical location, number of kids under 18, major occupation, tenure, tenure square and husband's yearly earnings. Industry and race are not statistically significant, so they are dropped out of the regression.

Table 4: Endogeneity Tests of Fulltime Wife's HI status

	Wu-Hausman Test		Durbin-Wu-Hausman Test	
	F (1,1258)	Prob>chi2	Chi2(1)	Prob>chi2
F-size IV	2.19	0.14	2.24	0.13
HHI IV	1.69	0.19	1.73	0.19
F-size, HHI	2.57	0.11	2.63	0.11

Notes:

1. The null hypothesis for Wu-Hausman test and Durbin-Wu-Hausman test is that wife's HI status is exogenous.
2. This "Durbin-Wu-Hausman" (DWH) test is numerically equivalent to the standard "Hausman test".
3. The regressors except HI in the wage equation are age, age square, education, geographical location, number of kids under 18, major occupation, tenure, tenure square and husband's yearly earnings. Industry and race are not statistically significant, so they are dropped out of the regression.

Table 5: Comparison between With and Without Indirect Effects in the Regression

	No Interactions		With Interactions			
	OLS	IV (GMM)	OLS		IV (GMM)	
	HI-Coef.	HI-Coef.	HI-Coef.	F-stat.	HI-Coef.	F-stat.
No IV	0.09 (0.002)		-0.06 (0.92)	1.05 (0.40)		
F-size IV		-0.22 (0.32)			-1.82 (0.25)	1.28 (0.15)
HHI IV		-0.06 (0.62)			-1.09 (0.44)	1.52 (0.04)

Notes:

1. When adding the interaction terms, the F-statistics is testing the null hypothesis that all the interaction terms equal to 0.
2. The p-values are in brackets.
3. The regressors except HI in the wage equation are age, age square, education, geographical location, number of kids under 18, major occupation, tenure, tenure square and husband's yearly earnings. Industry and race are not statistically significant, so they are dropped out of the regression.

Table 6: Wage differentials at different quantiles for different models

		Total wages and wage differentials			Wages explained by Constant and residual			Wages explained by observable characteristics ($x\beta$) and Oaxaca decomposition					
		W_{hi}	W_{nohi}	$\Delta W_{hi-nohi}$	W_{hi}	W_{nohi}	$\Delta W_{hi-nohi}$	W_{hi}	W_{nohi}	$W_{nohi/hi}$	$\Delta W_{hi-nohi}$	$\Delta W_{hi-nohi/hi}$	$\Delta W_{nohi/hi-nohi}$
No instrument	Mean	2.77	2.51	0.26	1.10	1.12	-0.02	1.66	1.36	1.51	0.30	0.15	0.15
	P10	2.08	1.73	0.35	0.56	0.61	-0.04	1.23	0.82	1.08	0.41	0.15	0.26
	P25	2.43	2.11	0.32	0.82	0.84	-0.02	1.43	1.10	1.28	0.33	0.15	0.18
	P50	2.77	2.53	0.24	1.11	1.13	-0.02	1.65	1.32	1.51	0.33	0.14	0.19
	P75	3.16	2.91	0.25	1.40	1.39	0.01	1.92	1.62	1.75	0.30	0.17	0.13
	P90	3.43	3.18	0.25	1.62	1.66	-0.04	2.12	1.86	2.02	0.26	0.10	0.16
Husband's firm size iv	Mean	2.59	2.71	-0.12	1.10	1.34	-0.24	1.48	1.39	1.59	0.09	-0.11	0.20
	P10	1.80	1.97	-0.17	0.57	0.79	-0.22	0.96	0.90	1.13	0.06	-0.17	0.23
	P25	2.26	2.36	-0.10	0.81	1.08	-0.27	1.23	1.14	1.32	0.09	-0.09	0.18
	P50	2.54	2.73	-0.19	1.10	1.35	-0.25	1.44	1.37	1.53	0.07	-0.09	0.16
	P75	3.00	3.07	-0.07	1.38	1.66	-0.28	1.72	1.64	1.86	0.08	-0.06	0.22
	P90	3.36	3.37	-0.01	1.69	1.88	-0.19	2.05	1.83	2.16	0.22	-0.11	0.33
Husband's HI status iv	Mean	2.61	2.69	-0.07	0.86	1.37	-0.51	1.74	1.27	1.77	0.47	-0.03	0.50
	P10	1.90	1.94	-0.04	0.38	0.81	-0.43	1.20	0.83	1.31	0.37	-0.11	0.48
	P25	2.26	2.36	-0.10	0.62	1.05	-0.43	1.51	1.01	1.51	0.50	0.00	0.50
	P50	2.65	2.69	-0.04	0.88	1.38	-0.50	1.73	1.22	1.75	0.51	-0.02	0.53
	P75	3.01	3.06	-0.05	1.10	1.69	-0.59	1.99	1.52	2.02	0.47	-0.03	0.50
	P90	3.33	3.39	-0.06	1.33	1.97	-0.64	2.28	1.81	2.28	0.47	0.00	0.47

Notes:

1. W_{hi} : wages for HI group. W_{nohi} : wages for NOHI group. $\Delta W_{hi-nohi}$: wage differentials between HI and NOHI groups. $W_{nohi/hi}$: wages for NOHI group based on HI group coefficients. $\Delta W_{hi-nohi/hi}$: wage differentials due to characteristic differences. $\Delta W_{nohi/hi-nohi}$: wage differentials due to compensation rate differences.
2. The regressors except HI in the wage equation are age, age square, education, geographical location, number of kids under 18, major occupation, tenure, tenure square and husband's yearly earnings. Industry and race are not statistically significant, so they are dropped out of the regression.

Table 7: Stochastic Dominance Test Results (Bootstrap 1000)

		FSD						SSD						
		d_{1MAX}	d_{2MAX}	d	$\Pr(d_1^* \leq 0)$	$\Pr(d_2^* \leq 0)$	$\Pr(d^* \leq 0)$	s_{1MAX}	s_{2MAX}	s	$\Pr(s_1^* \leq 0)$	$\Pr(s_2^* \leq 0)$	$\Pr(s^* \leq 0)$	
No instrument	Total wage	-0.012	0.197	-0.208	0.741	0.000	0.741	-0.012	51.982	-0.082	0.915	0.000	0.915	
	Constant + residual	0.038	0.022	0.390	0.000	0.024	0.024	6.102	0.023	0.405	0.088	0.366	0.454	
	$(X\hat{\beta})$ and Oaxaca decomposition	$X\hat{\beta}$ gap	-0.027	0.322	-0.479	0.991	0.000	0.991	-0.047	89.248	-0.827	1.000	0.000	1.000
		X gap	-0.010	0.172	-0.181	0.724	0.000	0.724	-0.036	47.704	-0.641	0.990	0.000	0.990
	$\hat{\beta}$ gap	-0.005	0.184	-0.090	0.595	0.000	0.595	-0.016	48.705	-0.267	0.929	0.000	0.929	
Husband's firm size iv	Total wage	0.141	0.017	0.293	0.000	0.036	0.036	24.085	-0.013	-0.217	0.000	0.873	0.973	
	Constant + residual	0.252	-0.020	-0.355	0.000	0.932	0.932	58.620	-0.026	-0.450	0.000	0.993	0.993	
	$(X\hat{\beta})$ and Oaxaca decomposition	$X\hat{\beta}$ gap	-0.001	0.146	-0.014	0.396	0.000	0.396	-0.016	36.612	-0.284	0.755	0.000	0.755
		X gap	0.123	-0.001	-0.020	0.000	0.384	0.384	30.163	-0.020	-0.344	0.000	0.981	0.981
	$\hat{\beta}$ gap	-0.024	0.241	-0.489	0.995	0.000	0.995	-0.024	68.486	-0.489	0.996	0.000	0.996	
Husband's HI status iv	Total wage	0.083	-0.000	-0.007	0.000	0.098	0.098	13.362	-0.004	-0.073	0.020	0.558	0.558	
	Constant + residual	0.470	-0.029	-0.490	0.000	0.999	0.999	111.655	-0.031	-0.515	0.000	1.000	1.000	
	$(X\hat{\beta})$ and Oaxaca decomposition	$X\hat{\beta}$ gap	-0.027	0.460	-0.455	1.000	0.000	1.000	-0.027	116.945	-0.455	1.000	0.000	1.000
		X gap	0.048	0.017	0.280	0.000	0.044	0.044	8.701	-0.023	-0.385	0.010	0.710	0.710
	$\hat{\beta}$ gap	-0.028	0.459	-0.582	0.999	0.000	0.999	-0.036	121.964	-0.749	1.000	0.000	1.000	

Notes:

- $X\hat{\beta}$ is the total gap due to characteristic differences; X is the characteristic gap; $\hat{\beta}$ is the compensation gap. Except for $\hat{\beta}$ gaps, the subscript 1 in d_{1MAX} and s_{1MAX} stands for the differences of HI minus NOHI; the subscript 2 in d_{2MAX} and s_{2MAX} stands for the differences of NOHI minus HI. For $\hat{\beta}$ gaps, the subscript 1 stands for the differences of NOHI/HI minus NOHI; the subscript 2 stands for the differences of NOHI minus NOHI/HI.
- The regressors except HI in the wage equation are age, age square, education, geographical location, number of kids under 18, major occupation, tenure, tenure square and husband's yearly earnings. Industry and race are not statistically significant, so they are dropped out of the regression.