

An Analysis of Income Differentials by Marital Status

▪ REGINA MADALOZZO*

RESUMO

O casamento não oficializado, coabitação, tem se tornado cada vez mais freqüente nas últimas décadas. O objetivo deste trabalho é examinar a relação entre os salários das mulheres casadas e das solteiras ou coabitantes. A literatura a este respeito mostra que, enquanto o prêmio financeiro para o casamento é verificado em diversos estudos e países quando o objeto de estudo são os homens, o resultado para mulheres não é conclusivo. A principal inovação do presente estudo é a existência de controles para seleção, tanto na escolha em participar da força de trabalho como de alterar seu estado civil. Regressões “switching” e decomposição de Oaxaca mostram a existência de uma penalização financeira para mulheres casadas. Corrigindo para ambos os tipos de seleção, a diferença nos salários das mulheres casadas com relação às coabitantes varia entre 49% e 53%, favorecendo as coabitantes. Este resultado aponta para a existência de uma penalidade ao casamento.

PALAVRAS-CHAVE

prêmio para casamento, salários, diferencial entre mulheres

ABSTRACT

Unmarried cohabitation has become a more frequently observed phenomenon over the last three decades, and not only in the United States. The objective of this work is to examine income differentials between married women and those who remain single or cohabit. The empirical literature shows that, while the marriage premium is verified in different studies for men, the result for women is not conclusive. The main innovation of my study is the existence of controls for selection. In this study, we have two sources of selectivity: into the labor force and into a marital status category. The switching regressions and the Oaxaca decomposition results demonstrate the existence of a significant penalty for marriage. Correcting for both types of selection, the difference in wages varies between 49% and 53%, when married women are compared with cohabiting ones, and favors non-married women. This result points to the existence of a marriage penalty.

KEYWORDS

marriage premium, wage, women differential

JEL CLASSIFICATION

J12, J16, J31

+ I am grateful to the seminar participants at the University of Illinois at Urbana-Champaign, Ibmec São Paulo and Brazilian Econometric Society 2002, and to Kevin Hallock, John Johnson IV, Roger Koenker, Anil Bera, Werner Baer, Marcelo Moura, Sherrilyn Billger and the two anonymous referees for helpful comments. Financial support from Capes-Brazil and CNPq Productivity Research Fellowship number PQ 307513/2007-6 are acknowledged.

* Ibmec – São Paulo. Endereço para contato: Rua Quatá, 300, 4º andar, São Paulo, SP. CEP 04546-042. E-mail: reginam@isp.edu.br.

(Recebido em julho de 2005. Aceito para publicação em maio de 2007).

Unmarried cohabitation has become a more frequently observed phenomenon over the last three decades, and not only in the United States. Recent legal changes in some countries give cohabitators the same legal standing as married couples.¹ In some countries, cohabitators sign affidavits, which stipulate that the union becomes a legal marriage after a pre-determined length of time.² In the United States, limitations on welfare receipt provide incentives for cohabitation as compared to marriage, because welfare benefits can be lost upon getting married. However, it is not only to escape from restrictive welfare rules that more people are choosing not to marry. Many young, well-educated persons opt to live together without being married. These persons are the so-called cohabitators.³ There was an increase in cohabitation in the United States over the last 20 years. The percentage of opposite sex couples cohabiting doubled from 1980 to 1999.⁴

The objective of this work is to examine income differentials between married women and those who remain single or cohabit. The empirical literature shows that, while the marriage premium is verified in different studies for men,⁵ the result for women is not conclusive. Because marriage has a strong connection with children, and women usually are the parent responsible for taking care of children, most studies consider children's impact on female wages. The earnings variation among women can fluctuate with the number of children (MOORE; WILSON, 1982) or with marital status (HILL, 1979). Concerning the number of children, women's wages present an increasing *family status penalty* in the 1980s if accounted for by the presence of children (WALDFOGEL, 1997). However, for those women who do not give up working at childbirth, there is no wage gap when compared to childless females (JOSHI; PACI; WALDFOGEL, 1999).

The main innovation of my study is the existence of controls for selection.⁶ Selection can occur when inclusion in a sample and presence of the variable of interest are both determined by the same unobservable factors. Not accounting for this problem causes

1 For instance, in Brazil, since the late 1990s, any couple living together for more than one year has the same rights and obligations to each other as they would have if they were legally married (Brazilian Law Code number 9278, May, 10th. 1996).

2 In Kenya, for instance, it is possible to avoid the costs of marriage, which are not only the expenses of the ceremony but also the bride wealth that is still common, using this maneuver. See Kabeberi-Macharia and Nyamu (1998).

3 The cohabitation trend is the subject of research across sociology, demography, and economics. For examples, see Brien, Lillard and Waite (1999), Bumpass and Sweet (1989), Carlson and Danziger (1999), Lillard, Brien and Waite (1995), Willis and Michael (1994) and Waite (1995).

4 See US Census Bureau at <http://www.census.gov> or the volumes of Statistical Abstracts of United States cited at the references.

5 See Allegretto and Arthur (1999) and Korenman and Neumark (1991).

6 Harkness and Waldfogel (1999) examine the differences in wage structure for females and control for selection into the labor force using Heckman's (1979) procedure. However, adequate controls for the selection problem is not the main point of their paper; neither is the interpretation of effects of children on mothers' wages. This happens because they cannot properly estimate the wage equation in the absence of important explanatory variables, such as experience and employee characteristics.

serious bias in an analysis.⁷ In this study, we have two sources of selectivity: into the labor force and into a marital status category. Women seem to have more flexibility in choosing to work, relative to men. The first selection control aims to solve this bias in the analysis. The second selection problem is the difference among women who make distinct choices about living arrangements, i.e., marrying, cohabiting, or remaining single.

This study supplements the literature on income differentials among cohabiting, single and married women by making use of the *switching regressions* as well the Oaxaca (1973)/ Kuhn (1987) decomposition method, as applied in Billger (2000). The switching regressions and the Oaxaca decomposition results demonstrate the existence of a significant penalty for marriage. When correcting only for selection into the labor market, the wage gap between married and cohabiting women favors married women. The predicted *marriage premium* is roughly 5%, using the switching regressions method. Using the same methods for the sample of single and married women, the *marriage premium* is 10%. However, this result is biased because it does not account for selection into marital status. Correcting for both types of selection, the difference in wages varies between 49% and 53%, depending on the method used, when married women are compared with cohabiting ones, and favors non-married women. This result points to the existence of a *marriage penalty*. Moreover, this *marriage penalty* can oscillate between 26% and 34% when the comparison is made between married and single females.

This paper is organized as follows: the next section describes the data and presents a demographic analysis of the sub-samples. Section 2 contains a discussion of the econometric methods to be used in the study and presents the results and their interpretation. Section 3 concludes and includes directions for future work.

1 DATA AND DESCRIPTIVE ANALYSIS

The data used in this paper come from the Annual Demographic Files of the Current Population Survey (CPS), the Annual Demographic Files for 1995, 1997, and 1999.⁸ The sample utilized in this study has 81,979 observations among married, cohabiting, and single women between the ages of 20 and 64 over these years.

One of the greatest advantages of using CPS data, besides its characteristic of being a very representative sample of the US population, is the possibility of identifying coha-

7 See more about the selection problem in Vella (1998).

8 The CPS uses the same sample for 4 consecutive months, keeping this sample out for the following 8 months and re-interviewing them for another 4 months. This process is called *rotation 4-8-4*.

bitors.⁹ Since 1990, the US Census has set apart unmarried partners, i.e., cohabitators, from spouses or housemates. The CPS began the cohabitators' identification in 1994.¹⁰ However, only the head of the household's partner is identified as a cohabitor by the CPS. The head of the household, or the "householder" before 1980, is the person who owns or rents the house unit, which presumably implies the person with higher income in the family.¹¹ In order to capture the *heads* who are cohabitators without misrepresenting the results,¹² it was necessary to implement another method. Using the household identification number, it was possible to link each husband and wife, as well as to connect unmarried partners as a couple, and to capture single women who are the family reference person.

Table 1 presents some descriptive statistics for the sample to be used in this study, classified as married, cohabiting, or single and using the CPS final weight for each individual. This table shows that cohabiting and single women are younger than married women and have fewer children.

The pattern on income data indicates that single women have higher wages and salaries (\$18,701), followed by cohabitators (\$16,607). However, when considering only those women with positive income, married women on average have larger wages and salaries than cohabitators (\$22,419 and \$20,281, respectively).

Even then, single women earn 4% more than married females. Hourly wage is a constructed variable. It is calculated dividing the variable "annual wages and salaries" by "usual number of hours worked per week" multiplied by "number of weeks worked last year".

9 The Panel Study of Income Dynamics (PSID) is another dataset that distinguishes cohabitators from married persons. However, this distinction is made only in the first year that the unmarried partner entered the family. After this, the cohabitor is treated as a spouse even if the couple did not actually marry, i.e. it is not possible to distinguish married from unmarried couples.

10 Until this date, the way to investigate cohabitation using the CPS data was to use the method called POSSLQ (Partners of the Opposite Sex Sharing Living Quarters). This methodology had the caveat of including only unmarried partners without children in the sample and, sometimes, accounting roommates as partners. The POSSLQ consists in identifying all households with exactly two adults of opposite sex who are unrelated. The Adjusted POSSLQ allows the inclusion of children, but it still does not capture all cohabitators and also includes some roommates as partners (CASPER; COHEN, 2000). The option of using the period after 1994 for the present study is justified by the inclusion of the correct sample of cohabitators.

11 If no such person exists, then any adult member, excluding roomers, boarders or paid employees can be characterized as the householder. If the married couple jointly owns or rents the house, then the householder can be either one.

12 Roughly 50% of the female cohabitators' sample is the head of the household, as can be seen in Table 1. Identifying the female as the *head* is either a peculiar characteristic of unmarried couples or means that, in these cases, the financial responsibilities are more equally shared between the partners.

TABLE 1 - DEMOGRAPHICS, CPS MARCH FILE (1995, 1997 AND 1999)
RESTRICTED TO WOMEN WITH AGE BETWEEN 20 AND 64, IN-
CLUSIVE

	Married	Cohabitor	Single
Age	41.82 (.042)	33.60 (.154)	33.80 (.106)
Annual Wages and Salaries	15,546 (82.89)	16,607 (294.67)	18,701 (219.78)
Annual Wages and Salaries (only for women with positive income)	22,419 (104.63) [47,365]	20,281 (329.29) [3,552]	23,227 (246.12) [7,875]
Hourly Wages	18.64 (.315)	16.07 (.924)	19.42 (.857)
Hourly Wages (only for women with positive wage)	19.69 (.319) [27,087]	16.47 (.945) [1,873]	19.82 (.875) [4,176]
Number of children	.34 (.003)	.19 (.008)	.26 (.006)
Employment (%):			
Part-time	20.61	15.50	14.48
Full-time	53.41	68.57	67.70
Unemployed	2.20	5.26	5.22
Others	23.78	10.67	12.60
Education(%):			
Less than high school	11.37	14.33	14.14
High school diploma	35.67	36.55	26.28
Some college	18.44	23.20	21.42
College degree	27.09	21.26	29.08
More than college	7.43	4.66	9.08
Status in the household (%):			
Head	21.14	45.24	100
Wife	78.86	-	-
Partner (cohabitor)	-	54.76	-
Race (%):			
White	79.07	74.06	55.81
Black	7.15	12.36	30.85
Hispanic	9.42	10.11	9.68
Other	4.36	3.47	3.66
# of Observations	67,887	4,335	9,757

Notes: 1) Standard errors are in parenthesis. 2) Number of observations in squared-brackets.

Concerning hourly wages, single women appear to have higher wages (19.42) than cohabiting (16.07) and married women (18.64). This pattern is maintained when observing only those women with positive income. Single women earn approximately the same hourly wages as married women (19.82 and 19.69, respectively), and both earn higher hourly wages than cohabiting women (16.47). A higher proportion of married women are engaged in part-time jobs (21%) than cohabiting (16%) or single women (15%). Cohabiting females are mostly allocated in full-time jobs (69%). Relative to education, a larger fraction of single women have a college degree or higher (38%) than married or cohabiting females (35% and 26%, respectively).

One common question in studies that compare different marital status is how to categorize cohabitators. Do they look more like married persons or singles? While cohabitators and single women have approximately the same work profiles, and both are more engaged in the labor force when compared to married women, they are very dissimilar concerning education, race, and hourly wages. Therefore, given these dissimilarities, it is reasonable to analyze married versus cohabiting women and married versus single women separately. For the rest of the analysis, the total sample is divided into these two sub-samples. This division facilitates the investigation of the direct effect of cohabitation or singlehood on income with respect to married women.

2 THEORETICAL FRAMEWORK AND EMPIRICAL RESULTS

The first step in this study is to analyze income differentials for women through a baseline wage regression.¹³ In this specification, the natural logarithm of hourly wages (in dollars) is explained by personal characteristics of the individuals and indicators for being an unmarried partner or single. The basic regression has as explanatory variables: experience,¹⁴ experience-squared, number of children younger than 6 years old, indicators for educational attainment,¹⁵ race,¹⁶ 2-digit industry, 1-digit occupation,¹⁷ residence in a metropolitan area and yearly indicators¹⁸ to account for the structure's

13 An unadjusted wage regression for women, even knowing the selection problem for this sample, is estimated. The reason for this is that the results from this estimation will serve as a comparison with the results of studies that deal or do not deal with the selection problems. The difference between this baseline regression and the adjusted one will be presented at the end of this paper.

14 This is a created variable: experience is the age minus the number of years of study minus six.

15 They are: less than high school, high school degree, some college, college degree, and post-college degree. Excluded category: less than high school.

16 They are: White, Black and Hispanic. Excluded category: White. Notice that the category white includes other minorities not specified as Black or Hispanic. The ethnicity variable that allows the identification of Hispanics do not allow the same for other minorities. Therefore, other minorities were included with the Whites.

17 All the regressions were also estimated using 2-digit occupation indicators. Results for these latter regressions point to the same direction as the ones presented here. However, the switching regressions and Oaxaca-Kuhn effects estimated by using 2-digit occupation indicators are abnormally higher than the ones estimated using 1-digit occupation indicators.

18 They are: 1995, 1997 and 1999. Excluded category: 1995.

difference over the years. For each sub-sample,¹⁹ an indicator for marital status was included. In the sub-sample married versus cohabitators, an indicator that assumes value 1 if the woman is a cohabitor and zero otherwise was included. In the sub-sample married versus single, an indicator that assumes value 1 if the woman is single and zero otherwise is the alternative.

Table 2 presents the results for this baseline regression. Column (1) uses the sample composed of married and cohabiting women and, column (2), the sample of married and single women. In all the regressions, robust standard errors were estimated to control for heteroskedasticity, as described in White (1980). All regressions are weighted by the final CPS weight. In both regressions, the return to experience is positive with decreasing marginal returns, as is clear from the negative sign on the variable experience-squared. The presence of young children has a negative influence on the income of their mothers. Returns to education are increasing, as expected. The coefficient on the indicator for Hispanics in both regressions is negative and significant. The coefficient for the Black indicator is not significant in both regressions.

The focus of this paper is on the impact of marital status on wages. Both the indicators for being a cohabitor and for being single have no significant estimated coefficients relative to being married. These results imply similar wages for married women when compared to cohabitators or singles. These results are interesting and suggest no income differential for women in distinct marital status categories. However, the sample is composed of only women and an extensive literature on the selection problem into the labor force shows that women have a different pattern of choosing to work than men. Besides the discontinuous participation in the labor force caused by childbearing,²⁰ it is necessary to account for the possibility that differences in unobserved characteristics, that will henceforth be called *ability*, influence female participation in the labor market and, consequently, the wage level.

Another source of selectivity in this study is selection into marital status. Given different characteristics, women opt to remain single, cohabit, or get married. This choice among different marital statuses can also reflect the sub-samples' unobservable characteristics. If these unobservable characteristics act both in the choice over marital status and the wage level, not accounting for these influences would bias the inference.

19 I.e., married versus cohabiting, and married versus singles.

20 Blau and Ehrenberg (1997) study the female role in the labor market. Goldin (1990) presents valuable research on the evolution and trend of female participation in the labor force.

TABLE 2 - BASIC REGRESSIONS – DEPENDENT VARIABLE: LN(HOURLY WAGE)

	Married and Cohabitors	Married and Singles
	(1)	(2)
Cohabitation indicator	.007 (.028)	-
Indicator for singles	-	-.009 (.021)
Experience	.031*** (.002)	.031*** (.002)
Experience squared	-.006*** (.0001)	-.001*** (.0001)
Children < 6	-.090*** (.014)	-.113*** (.014)
High school	.199*** (.031)	.224*** (.031)
Some college	.247*** (.034)	.297*** (.034)
College	.455*** (.035)	.502*** (.035)
Pos-college	.704*** (.041)	.733*** (.040)
Black	.048* (.026)	-.027 (.023)
Hispanic	-.074*** (.027)	-.074*** (.027)
Industry indicators	Yes	Yes
Occupation indicators	Yes	Yes
Metropolitan Area Indicator	Yes	Yes
Year indicator	Yes	Yes
Constant	1.57*** (.323)	1.59*** (.318)
Adjusted R-squared	.17	.17
# of Observations	27,087	29,390

Notes: 1) Robust standard error are in parenthesis. 2) Statistical significance by: * at 10%, ** at 5% and *** at 1%.

The technique used to deal with selection into the labor force is based on Heckman (1979). For the second selectivity problem, related to the choice of marital status, the switching regressions technique is used, as discussed in Maddala (1986). Notice that not accounting for these selectivity problems would result in biased estimates.

2.1 Selection Into Labor Force Participation

One way to deal with selectivity into the labor force, as noted previously, is to use the Heckman (1979) model. Women with different characteristics or *abilities* choose to engage or not engage in the labor force. While the presence of young children has the effect of increasing the cost of working, educational attainment and experience can have the opposite influence. Therefore, to estimate the probability of participating in the labor force, we use the number of children younger than 6 years old,²¹ indicators for education, and age and age-squared. Prior studies show that non-white women have a stronger commitment to the labor market. To account for this effect, indicators for race were included. Besides these intuitive variables, indicators for marital status were also included.²² The expected sign for both indicators is positive, given that the household production theory affirms that division of work is efficient when each member of a family dedicates their time to the more productive job.²³ Men usually receive relatively better compensation for their time in the labor market than in home production. Thus, the expectation is that married women dedicate more time to *home* tasks and less to the labor market, and this would imply a different probability of working given the marital status choice.

Table 3 presents the results for the estimation of the probit regression, which is specified as:

$$\Pr(\text{working}) = \Phi[\eta_1 + \eta_2 \text{age} + \eta_3 \text{agesq} + \eta_4 \text{child} + \sum_i \delta_i \text{educ}_i + \sum_j \zeta_j \text{race}_j + \kappa D + \varepsilon] \quad (1)$$

Where $\Phi(\cdot)$ is the cumulative standard normal distribution, *educ* accounts for the four indicators on education attainment, and *race* represents the two race indicators. *D* is the indicator for marital status, as explained before. The results are very similar for both samples as shown in columns (1) and (2).

21 In order to check for robustness, alternative specifications were tried, like the inclusion of number of children less than 18 years old. Results varied slightly in magnitude, but the conclusions remain the same. Also checking robustness, an alternative concept of the dependent variable was used. Instead of using *wages and salaries*, the tested alternative was *income from the longest job*. Final results remain basically the same. However, this latter specification is not the most sensible way to analyze a wage differential. Therefore, only the results from the regressions with cited specifications in the main text are reported.

22 Different indicators were included depending on the sample composition. For instance, in the sample with married and cohabiting women, an indicator for cohabitators was included. In the sample of married and single women, an indicator for singles was included as a regressor.

23 See Becker (1973) and Angrist and Evans (1998).

TABLE 3 - PROBIT RESULTS FOR PROBABILITY OF BEING IN THE LABOR FORCE⁽¹⁾

	Married and Cohabitors (1)	Married and Singles (2)
Cohabitation indicators	.181*** (.023)	-
Indicator for singles	-	.172*** (.016)
Age	.090*** (.003)	.086*** (.003)
Age squared	-.001*** (.0001)	-.001*** (.0001)
Children < 6	-.381*** (.009)	-.384*** (.008)
High school	.488*** (.016)	.500*** (.016)
Some college	.590*** (.018)	.605*** (.017)
College	.751*** (.017)	.780*** (.017)
Pos-college	1.01*** (.025)	1.04*** (.024)
Black	.131*** (.022)	.035* (.019)
Hispanic	-.138*** (.015)	-.140*** (.015)***
Constant	-1.12*** (.071)	-1.03 (.068)
Pseudo R-squared	.09	.10
# of observations ⁽²⁾	77,913	83,845

Notes: 1) Robust standard error are in parenthesis. 2) Statistical significance by: * at 10%, ** at 5% and *** at 1%. 3) The number of observations on this table may differ from the one on Table 2, since the latter is a weighted estimation over the sample.

As expected, age and age-squared account for the concave experience effect: the former is positive and the latter negative. Both are significant. Young children have a negative influence on participation in the labor force. Consistent with the expected positive and increasing returns to education, the higher the educational degree attained, the larger the probability of working. For the sample of married women and cohabitators, there is a positive and significant effect of being black, implying a higher probability of participation into the labor force by Black women. In column (2), relative to the sample of married and single women, the same coefficient is not significant. Hispanic accounts for a lower probability of participation in the labor

force, and it is significant in both samples. In addition, as anticipated, the effect of not being married, i.e. being a cohabitor or single, on the probability of working is positive and significant. The correction for selection bias for participating in the labor force is attained by the Inverse Mills Ratio (IMR) as one of the regressors in the wage equation. The IMR is defined as:

$$IMR = \frac{\varphi(Z'\gamma / \sigma_0)}{\Phi(Z'\gamma / \sigma_0)} \quad (2)$$

Where $Z'\gamma$ represent the regressors in the probit equation, and σ_0 is the standard deviation of the residuals from the original equation that had the sample selection problem.

Table 4 presents the results for both samples for the logarithmic wage regression augmented with the additional regressor (2), as shown in Equation (3):

$$\ln(wage) = \sum_j \beta_j X_j + \gamma D + \delta IMR + \varepsilon \quad (3)$$

Where X_j represents the covariates as described previously,²⁴ D is the dummy for the marital status, and IMR is the variable that controls for selection into the labor force. A note of attention must be made at this point. While on Table 2 the estimated coefficients indicate the marginal effect of each independent variable on the dependent one, the results presented in Table 4 and the following ones could not be interpreted in the same way. When we use the Heckman correction procedure – the inclusion of IMR on the regression –, the marginal effect of each individual variable on the dependent one must take in account the indirect effect of the IMR coefficient.^{25,26} Columns (1) and (2) present the results when controlling for participation in the labor force by the sub-samples of married and cohabiting women or married and singles, respectively. Notice that, in order to have an identified model, some authors suggest that the probit regression, Equation (1), should include at least one variable that is not a regressor in Equation (3). Therefore, from the wage regressions, the variable *children < 6* was excluded. It also makes sense to proceed with this exclusion because the number of children is more likely to influence participation into the labor market than the wage level of the mother. In addition, the variable *age squared*, which is not included in the wage regression, enters in the probit model. When included in a wage regression, the variable *age squared* or *experience squared* accounts for the existence of

24 They are: experience, experience-squared, indicators for educational attainment, race, 2-digit industry, 1-digit occupation, residence in a metropolitan area and yearly indicators.

25 More about this on Hoffmann and Kassouf (2005) and Saha *et al* (1997).

26 For the final calculations of the differences between married and cohabitators or single women, we use this indirect effect, as the reader would be able to check on Section II.

a concavity in the wage.²⁷ The same reasoning can be used here. The probability of participating in the labor force may increase with time and after some determined age this probability decreases. The significance of this coefficient estimative reinforces this interpretation.

TABLE 4 - REGRESSION RESULTS OF LN(HOURLY WAGE), CONTROLLED BY MRP (WORK)

	Married and Cohabitors	Married and Singles
	(1)	(2)
Indicator for marital status	-.043 (.028)	-.056** (.022)
Mills Ratio Prob(Working)	-.621*** (.069)	-.693*** (.069)
Experience	.021*** (.003)	.021*** (.002)
Experience squared	-.001*** (.0001)	-.001*** (.00001)
High School	.062* (.035)	.064* (.035)
Some college	.090** (.039)	.113*** (.039)
College	.261*** (.042)	.274*** (.042)
Pos-college	.464*** (.050)	.454*** (.049)
Constant	2.00*** (.327)	2.05*** (.323)
Adjusted R-squared	.17	.18
# of observations	27,087	29,390

There are caveats for both estimated regressions in columns (1) and (2). The first is the negative sign on the Inverse Mills Ratio for the Labor Force Participation. One should expect that the larger the unobservable characteristics that positively influence the participation of women into the labor force, the greater should be the expected income. However, the results point in the opposite direction. For these women, selectivity into the labor force exists, as the significance on the IMR coefficient indicates, although women who expect to be less well remunerated by their work actually are

²⁷ Age and experience are different variables. We opt to use the experience and experience squared in this work. Following the literature on this research area, the experience variable is a proxy for effective experience calculated by age – 6 – years of education. This procedure is widely used because most databases do not have the actual experience of individuals, only their age and education.

more likely to participate in the labor force. This unexpected result is not exclusive to this paper. Vella (1998), using NLS data, finds the same negative sign in the IMR coefficient. His result also implies that selectivity into the labor force works in a reverse way. Women with characteristics that predict receiving lower wages are more likely to participate in the labor market.

One possible explanation of this odd result is that women in these samples have a greater need to work. It is plausible to think that a single mother has no choice but to work, especially if she and her children depend entirely on her labor. The Welfare Reform Act of 1994 may have contributed to this result. By these reforms, no one could be a welfare recipient for more than 3 years. This change is only valid after 1999, however we could reasonably argue that this may have had some effect on the preceding years by adjustment to the future implementation. In addition, the partner or husband's income could influence participation in the labor force. Generally, unions are made in a similar income range. Therefore, poorer, less educated women in general would be paired with men with similar characteristics and might have no other choice but to work, independent of the smaller compensation for their own characteristics (BECKER, 1973).²⁸

Finally, the coefficients on the Table 2 for the central variables, cohabitating in column (1) and single in column (2), changed substantially once we account for selection into labor force participation, on Table 4.²⁹ This means that if we only correct for selectivity into the labor force using Heckman's procedure,³⁰ there is a *marriage premium* for the sample of married women compared to single ones. However, as discussed before, there is also selectivity in the marital status choice. The next step is to deal with this selectivity bias, which will be done by using the *switching regressions* approach.

2.2 Selection into different family categories

In this sub-section, results correct for selection into different marital statuses. The dataset has no information on the background of these women (e.g., parents' income and education) besides their education and race profile, but some of the more rele-

28 Neal (2001) has an interesting theoretical paper reasoning that economically disadvantaged women may choose to remain single when an economic crisis happens. He argues that the creation of a welfare system may have reinforced this decision, and also may have created the opportunity for these single women to have children.

29 Notice that the race indicators are not included on Table 4, because their estimations were not significant.

30 This result is reinforced when using *switching regressions*. Table 7 presents the results for both sub-samples, with a significant difference between wages paid to married and to unmarried women, the *marriage premium*, when correcting only for selection into the labor force. Sub-section 2.3 presents these results.

vant information that could influence the decision to get married are included in the empirical model.

Here it is necessary to highlight the importance of conducting this second selection correction. As when we estimated the wage regression correcting for the individual selection into the labor market in order to account for the effect of marriage status, we also have to correct for the individual selection into each family category. For instance, if the most able women choose to remain single by their own; while the less able choose to marry, and then not accounting of this type of selection would mistakenly include a kind of ability bias on the estimated coefficient of marriage status. The second selection correction will allow the correction for this bias and also the calculation of the real effect of marriage on wages.

First, by the demographics, one can conclude that age is important in this decision: married women are older than singles or cohabitators. The presence of younger children could influence this decision in either direction. On one hand, having a child may motivate mothers to marry because they could share the responsibilities of educating and sustaining their children. However, on the other hand, single mothers can lose part of their income when they marry. For instance, child support receipts can decrease if the single mother opts to marry.³¹

The education profile is included to explore the differences in the composition of these groups. Using the demographics in Table 1, it is evident that Black women are more likely to stay single. In the opposite direction, Hispanic women are more likely to marry. The inclusion of race indicators accounts for these differences.

The variable that is included in the probability model to cohabit or stays single, which is not included in the wage estimation, is the indicator for owning a house. Spending a considerable sum of money to buy a house can signal an inclination toward, or a readiness, for stability.³² As the papers that analyze male returns to marriage emphasize, one of the possible motives for the marriage premium could be that stability in one's personal life can result in on-the-job productivity. The expected sign on this variable is negative for both regressions, probability of cohabiting and probability of staying single, implying that owning a house, as one would suspect, is related to being married.

31 See Hu(1999) and Veum(1992).

32 Of course, buying a house also means that the person has some wealth. This fact does not contradict the fact that the person who buys a house is also the one who has more inclination toward stability. Even considering a house only as a sign of wealth, like an investment, we also could think that an investment is a sign of stability *per se*. Both ways, the variable *owning a house* can be considered a good proxy for marriage.

The probability model of cohabiting or being single can be expressed by the following equations, similar to Equation (1):

$$\Pr(\text{cohab} / \text{single}) = \Phi[\eta_1 + \eta_2 \text{age} + \eta_3 \text{child} + \eta_4 \text{housing} + \sum_i \delta_i \text{educ}_i + \sum_j \zeta_j \text{race}_j + \varepsilon] \quad (4)$$

Where *housing* is an indicator for owning a house. Table 5 presents the results for these regressions. As expected, age negatively influences the probability of being a cohabitor or single. Having a young child has the same effect, which suggests that the benefits of staying single and, possibly, maintaining alternatives sources of income (e.g., child support or alimony) are more than compensated for by the benefits of sharing the responsibilities of parenthood.

Schooling has different effects for the samples. Having more years of education implies a lower probability of entering into a cohabitation union, but this has the opposite consequence on the probability of staying single. The higher the education degree, the more likely a woman will stay single. As anticipated, the indicator variable for Black is positive and highly significant for probability to continue single. The Hispanic indicator is negative for the probability of entering into a cohabitation union, but smaller and positive for the probability of remaining single. Owning a house has the intuitive sign: it is more negative for the probability of staying single than for cohabiting. For both regressions, the estimated coefficients are highly significant.

The *switching regressions* method, described in Madalla (1986) and applied in Billger (2000), uses the results from Table 5 to control for *selection* into family status. This method is both an alternative to and an adjustment of the Heckman procedure for the cases where more than one type of selection is involved in the regressions. It fits in this study in order to control for the two types of selectivity that could bias the final analysis.

TABLE 5 - PROBIT RESULTS FOR PROBABILITY OF BEING A COHABITOR OR PROBABILITY OF BEING SINGLE

	Married and Cohabitors	Married and Singles
	(1)	(2)
Age	-.038*** (.001)	-.030*** (.001)
Children < 6	-.555*** (.019)	-.452*** (.013)
High school	-.087*** (.026)	-.134*** (.022)
Some college	-.074*** (.029)	.045* (.023)
College	-.243*** (.029)	.145*** (.022)
Pos-college	-.171*** (.041)	.452*** (.029)
Black	.147*** (.029)	.904*** (.019)
Hispanic	-.245*** (.026)	.030 (.020)
Own a house	-.558*** (.017)	-1.05*** (.014)
Constant	.564*** (.044)	.573*** (.035)
Adjusted R-squared	.15	.26
# of observations	77,913	83,845

Notes: Robust standard error are in parenthesis. Statistical significance by: * at 10%, ** at 5% and *** at 1%.

In the switching regressions, the sub-samples are divided again, now by each marital status: married, cohabitators and single women. The estimation proceeds by two wage equations:³³

$$y_1 = \ln(\text{wage}_{\text{cohabitators}}) = \sum_i \beta_{ci} X_i + \varepsilon_1 \quad (5)$$

$$y_2 = \ln(\text{wage}_{\text{married}}) = \sum_i \beta_{mi} X_i + \varepsilon_2 \quad (6)$$

33 In the text, the explanation is linked to the sample of married and cohabiting women. The same is applied to the sample of married and single women, by replacing cohabitators for singles.

The variable y_k (in this case, wage profile and $k=1,2$) is assumed to follow different probability laws for married and cohabiting women. There is also an Indicator Function, which takes value 1 when the selected characteristic is present (in this case, being a cohabitor) and zero otherwise. Making use of the Heckman (1979) selectivity theory, we have:

$$E(\varepsilon_1 | y_1 \geq y_2) = -\sigma_{1\varepsilon} \varphi(Z) / \Phi(Z) \quad (7)$$

where Z are possible explanatory variables for the occurrence of the selected characteristic,³⁴ which is being a cohabitor.

As a practical matter, the estimated equations are:

$$y_1 = \ln(\text{wage}_{\text{cohabitor}}) = \sum_i \beta_{ci} X_i + \sigma_{1\varepsilon} \left(\frac{\varphi(\gamma'Z)}{\Phi(\gamma'Z)} \right) + \varepsilon_1 \quad (8)$$

$$y_2 = \ln(\text{wage}_{\text{married}}) = \sum_i \beta_{mi} X_i + \sigma_{1\varepsilon} \left(\frac{\varphi(\gamma'Z)}{1 - \Phi(\gamma'Z)} \right) + \varepsilon_2 \quad (9)$$

The estimated coefficients for equations (8) and (9) will be useful in the next subsection. Results for these estimations are presented on Table 6 – all regressions include the indicators for occupation, industries, metropolitan area and year. Notice that the two IMR are included on the regressions, so the estimated coefficients must be interpreted using this correction to calculate the marginal effects. In that section, the results are used to estimate the predicted income differentials by marital status using the switching regressions model and the Oaxaca-Kuhn decomposition.³⁵

34 The other regressors are the same expressed in Table 5: age, number of children less than 6 years old, indicators for education and race, indicator for owning a house and a constant.

35 Tables for these estimations are available upon request.

TABLE 6 - RESULTS FOR SWITCHING REGRESSIONS – MARRIED AND COHABITORS – DEPENDENT VARIABLE: LN(HOURLY WAGE)

	Married			Cohabitors		
	(1)	(2)	(3)	(4)	(5)	(6)
Mills Ratio Prob(Cohabit)	-	-	1.08*** (.151)	-	-	.565*** (.165)
Mills Ratio Prob(Working)	-	-.599*** (.073)	-1.03*** (.096)	-	-1.26*** (.259)	-2.42*** (.418)
Experience	.034*** (.002)	.020*** (.003)	-.002 (.004)	.029*** (.008)	.010 (.008)	-.030*** (.014)
Experience squared	-.0006*** (.0001)	-.0002*** (.0001)	.0002*** (.00001)	-.0003 (.0002)	.0003 (.0002)	-.0009*** (.0003)
High School	.190*** (.033)	.047 (.037)	-.092** (.041)	.373*** (.094)	.078 (.115)	-.294* (.166)
Some College	.242*** (.036)	.072* (.042)	-.099** (.047)	.430*** (.104)	.063 (.131)	-.418** (.200)
College	.447*** (.037)	.241*** (.044)	.003 (.055)	.700*** (.109)	.250* (.147)	-.399 (.247)
Pos-college	.700*** (.043)	.448*** (.053)	.175*** (.065)	.918*** (.147)	.402*** (.185)	-.321 (.280)
Black	.048* (.027)	.037 (.027)	.035 (.027)	.001 (.085)	-.003 (.084)	-.003 (.084)
Hispanic	-.091*** (.028)	-.039 (.029)	-.043 (.029)	-.037 (.096)	.033 (.097)	.015 (.096)
Constant	1.50*** (.338)	1.99*** (.339)	2.69*** (.350)	1.65*** (.166)	2.45*** (.237)	2.54*** (.239)
R-squared	.16	.17	.17	.20	.21	.22
# of observations	25,214	25,214	25,214	1,873	1,873	1,873

	Married			Singles		
	(1)	(2)	(3)	(4)	(5)	(6)
Mills Ratio Prob(Single)	-	-	.545*** (.068)	-	-	.331*** (.063)
Mills Ratio Prob(Working)	-	-.589*** (.073)	-.858*** (.081)	-	-1.25*** (.220)	-1.72*** (.243)
Experience	.034*** (.002)	.021*** (.003)	.005* (.003)	.043*** (.005)	.026*** (.006)	.006 (.007)
Experience squared	-.0006*** (.0001)	-.0002** (.0001)	.0001 (.0001)	-.0008*** (.0001)	-.0002** (.0001)	.0001 (.0001)
High School	.190*** (.033)	.047 (.037)	-.049 (.039)	.450*** (.092)	.108 (.104)	-.090 (.113)
Some College	.242*** (.036)	.071* (.042)	-.031 (.043)	.668*** (.096)	.218* (.115)	.004 (.124)
College	.447*** (.037)	.239* (.045)	.119** (.047)	.899*** (.098)	.351*** (.129)	.119 (.138)
Pos-college	.700*** (.043)	.446*** (.053)	.335*** (.055)	1.01*** (.107)	.397*** (.143)	.174 (.151)
Black	.048* (.027)	.051* (.027)	.181*** (.030)	-.201*** (.047)	-.161*** (.047)	.021 (.058)
Hispanic	-.091*** (.028)	-.040 (.029)	-.005 (.029)	-.132* (.075)	-.040 (.076)	.017 (.077)
Constant	1.50*** (.338)	1.98*** (.339)	2.44*** (.333)	1.26*** (.112)	2.26*** (.207)	2.12*** (.194)
R-squared	.16	.17	.17	.24	.25	.25
# of observations	25,214	25,214	25,214	4,176	4,176	4,176

Notes: 1) Robust standard error are in parenthesis. 2) Statistical significance by: * at 10%, ** at 5% and *** at 1%.

2.3 - Predicted income differentials

In this analysis, two different methods will be applied. The first one uses the switching regressions procedure to calculate a percent wage differential, as in Billger (2000). The second technique is closest to the approach of Hallock, Hendricks and Broadbent(1998), who use an individual-based form of the Oaxaca decomposition as introduced in Kuhn (1987).

Using the estimated parameters calculated from equations (8) and (9), it is possible to predict wages for cohabiting women as if they were cohabitators or married. With the results from there regressions, the same calculations are used to predict wages for single women as if they were single or married. After these computations, the results are used to calculate the wage differentials for the different samples. The second method of analyzing income differentials is Kuhn's extension of the Oaxaca decomposition. The original work of Oaxaca (1973), on discrimination against women, suggested that the income differential measure should be:

$$D = \frac{\frac{W_m}{W_f} - \left(\frac{W_m}{W_f}\right)^0}{\left(\frac{W_m}{W_f}\right)^0} \quad (10)$$

where $\left(\frac{W_m}{W_f}\right)$ represents the observed male-female wage ratio and $\left(\frac{W_m}{W_f}\right)^0$ represents the male-female wage ratio without the existence of discrimination.

In order to get this measure, he suggested the use of the following regressions. Both male and female data, separately, would be regressed as shown in equation (11)

$$\ln(W_i) = \sum_i \beta_i X_i + u_i \quad (11)$$

where W_i is the hourly wage rate for the i -th worker, X_i represents a vector of individual characteristics, β is the regression's coefficients and u_i represents the error term.

Having the estimated $\hat{\beta}_j^m$ coefficients for the male sample and the estimated $\hat{\beta}_j^f$ for the female sample, we use them with the sample means, and it is possible to get the statistic \bar{D} :

$$\bar{D} = \sum_j \hat{\beta}_j^m \bar{X}_j^f - \ln(\bar{W}^f) \quad (12)$$

where \bar{X}_j^f represents the average of each variable that composes the vector of individual characteristics for females, and \bar{W}^f the average female wage. The discrimination factor is the difference between the observed mean of the female wage from the wage that women would have if they were evaluated as men and considering the observed characteristics of women.

Kuhn (1987) extends this derivation. Instead of using the average of each variable that composes the vector of individual characteristics of the sample, he suggests using the individual specific measures, with two alternative measures. They are:

$$\hat{D}_i^1 = \sum_j \hat{\beta}_j^m X_j^f - \ln(W_i^f) \quad (13)$$

$$\hat{D}_i^2 = \sum_j \hat{\beta}_j^m X_j^f - \sum_j \hat{\beta}_j^f X_j^f \quad (14)$$

\hat{D}_i^1 measures the income differentials over the actual wage of each woman and \hat{D}_i^2 uses an estimative of the wage for each woman. The choice between these two depends on the women's unobserved characteristics.³⁶ Neither one of these measures would be preferable to the other, unless there are assumptions about unobserved characteristics. In the present work, I use both equations (13) and (14) to analyze the data. Only the results from equation (14) are presented.³⁷

Tables 7 and 8 present the results of these two methods. Table 7 presents the results for the switching regressions and Table 8 for the Oaxaca-Kuhn's decomposition.³⁸ Notice that the IMR effects (on labor force participation and family categories) are included on these calculations.

36 Kuhn (1987) says that Equation (13) is preferable when the unmeasured ability is sector-specific, and Equation (14) is preferable when the unmeasured ability is general.

37 Table 7 presents the average of the difference between what married women were supposed to receive if they were cohabitators and what they were supposed to receive being married, using the estimated parameters from equation (11). The average is the same for both equations (13) and (14). Only the standard errors change. Equation (14) has smaller standard errors. However, even in the results from equation (13), which have bigger standard errors, the predicted coefficients for the difference between married and cohabitators (or singles) were still significant. Their t-statistics were between 10.9 and 42.3.

38 Notice that Table 8 presents the results for the Oaxaca-Kuhn predicted wage differences in percents.

TABLE 7 - WAGE DIFFERENTIALS USING SWITCHING REGRESSIONS

PANEL A: Predicted hourly wage differentials for cohabiting women					
	Predicted Wage Differentials [(2)-(3)]/(3)	Cohabiting Coefficients Mean Predicted Wage	Married Coefficients Mean Predicted Wage	Difference in Predicated Wages	t-statistic for Difference
	(1)	(2)	(3)	(4)	(5)
Baseline	.011	10.22	10.10	.12	4.25
Controlling for Labor Force Participation Selection	-.052	10.31	10.87	-.56	-21.42
Controlling for both PLF and Marital Status Selection	.528	10.85	7.10	3.75	94.03

PANEL B: Predicted hourly wage differentials for single women					
	Predicted Wage Differentials [(2)-(3)]/(3)	Single Coefficients Mean Predicted Wage	Married Coefficients Mean Predicted Wage	Difference in Predicated Wages	t-statistic for Difference
	(1)	(2)	(3)	(4)	(5)
Baseline	-.034	10.78	11.16	-.38	-13.94
Controlling for Labor Force Participation Selection	-.099	10.71	11.89	-1.18	-45.61
Controlling for both PLF and Marital Status Selection	.256	10.75	8.56	2.19	98.17

TABLE 8 - OAXACA-KUHN PREDICTIONS FOR COHABITING/SINGLE WOMEN SAMPLE (PREDICTED WAGE AS COHABITOR/SINGLE - PREDICTED WAGE AS MARRIED) ÷ PREDICTED WAGE AS MARRIED

	As Cohabitors	As Single
	(1)	(2)
Baseline	.045 (.004)	.032 (.003)
Controlling for Labor Force Participation Selection	-.031 (.004)	-.039 (.003)
Controlling for PLF and Marital Status Selection	.492 (.009)	.336 (.003)

Note: Standard errors are in parenthesis.

Panel A of Table 7 presents the results for the cohabitators sample, and Panel B shows the results for the sample composed of singles. Using the baseline specification, cohabiting women, keeping their own characteristics, receive hourly wages 1% higher than if they were married. When controlling for selection into the labor force, this difference goes in favor of married women. The marriage premium is 5.2%. However, as discussed previously, the comparison between married women and cohabitators (or singles) should account for the selectivity in the choice of marital status. Accounting for the second type of selectivity, cohabitators earn an income 53% higher than if they were married.

The results for singles are similar. Without controlling for any type of selectivity, i.e. using the estimated parameters from the baseline regression, single women earn annual wages and salaries 3.4% lower than if they were married. Controlling only for selectivity into the labor market, this difference is 9.9%, still in favor of married women. However, in the final specification, which accounts for both selectivity into the labor force and selectivity in the choice of marital status, shows that single women have incomes 25.6% higher than if they were married.

These results point to the existence of a *marriage penalty* for women. Ignoring both selectivity problems would bias the results. Correcting only for the selectivity into the labor force would bias them implying a *marriage premium* for both sub-samples. Only the specification that accounts for both selectivity problems gives the result of a wage premium for non-married women between 25.6% (singles' case) and 53% (cohabitators' case).

Table 8 presents the results for the measure of income differentials developed by Kuhn for the two sub-samples.³⁹ Column (1) shows the results of the predicted difference between what cohabiting women earn, with their own characteristics, being cohabitators and what they would be predicted to earn being married. Using the baseline specification, cohabiting women earn 4.5% more being cohabitators instead of being married. Controlling for the selection into the labor force, this difference flips to 3.1% in favor to married women. However, we should also control for the selectivity into marital status. Using the final specification, the difference between what cohabiting women would earn by being cohabitators and what they would earn if they are married is 49.2%, reinforcing the marriage penalty estimated by the switching regressions method.

Column (2) in Table 8 presents the estimated difference between what single women would earn by remaining single and what they would have earned if they were mar-

³⁹ Table 8 expresses the values for the estimation of:

$[(\text{wage as cohabitor/single}) - (\text{wage as married})] \div \text{wage as married}$

In order to get the percents values, it is necessary to calculate the exponential value of this difference.

ried. The baseline regression shows a difference of 3.2% between the two predictors. Controlling only for selection into the labor market pushes the difference to 3.9%, but in favor of married women. The assumed correct specification is on the last line, which controls for both selection problems (into the labor force and in the choice of marital status). This last specification predicts a difference of 33.6% between what single women is predicted to earn by remaining single and what they would have earned if they were married.

By two alternative methods, the switching regressions procedure and the Oaxaca-Kuhn decomposition, results indicate the existence of a marriage penalty for women, when adequately controlling for both selection problems. The magnitude of this penalty varies with the chosen procedure, however both are consistent on the direction of the difference, favoring cohabitators and single women.

3 CONCLUSION AND FUTURE DEVELOPMENTS

The main goal of this paper is to empirically investigate women's income differentials by marital status. The motivation for this relies on the fact that the marriage premium for males is a well-known result, but for females the existence of a penalty for being married or a premium for being single or cohabiting is a topic that has received much less attention. Using data from the CPS for 1995, 1997 and 1999, and controlling for two types of selectivity, using techniques as in Maddala (1986) and Oaxaca(1973), empirical results show that married women have lower pay than non-married or cohabiting women. My estimates indicate a statistically significant income gap between married and cohabiting women in the range of 49% to 53%. When comparing married with single women, this difference increases. Using the *switching regressions* method, the difference is 25.6% in favor of single women. By the Oaxaca-Kuhn decomposition, the difference of annual income between married and single women is 34%.

This paper shows that both selection into the labor force and in the choice of marital status matter. Not accounting for them would seriously bias the final analysis, even implying a non-existent marriage premium. Accounting for selection into the labor force is important, as other cited references indicate. However, not accounting for the choice of marital status would wrongly predict that income differences between married and cohabiting (or single) women favor the former. Controlling for both types of selection, we have a consistent result of the difference for each sub-sample and conclude that a marriage penalty exists for women. Because the income difference between women in distinct marital status categories has received little attention up to now, there are some avenues for further investigation. One of the possible future

developments of this research is to divide the sample by race. Because white women have a somewhat different profile for work than non-white women, this investigation could shed some light on the subject. Another potential development would be a theoretical model for the existence of income differentials among women given their marital status. Marriage, which may give men the appearance of stability or greater ability, may have an opposite significance for women. A model that appropriately explores these ideas would be interesting.

Finally, the study of cohabitation in a panel data sample would be appealing. It would be possible to analyze decisions over education, labor force participation, impacts on future generations, and to see what happens to individuals' labor market outcomes when their marital status changes. I hope that my work is a useful first-step in this area of research.

4 REFERENCES

- ALLEGRETTO, S.; ARTHUR, M. An empirical analysis of homosexual/heterosexual male earnings differentials: unmarried and unequal? *Industrial and Labor Relations Review*, forthcoming, 1999.
- ANGRIST, J.; EVANS, W. Children and their parents' labor supply: evidence from exogenous variation in family size. *American Economic Review*, v. 83, n. 3, p. 450-477, 1998.
- BAUMAN, Kurt J. Shifting family definitions: the effect of cohabitation and other nonfamily household relationships on measures of poverty. *Demography*, v. 36, n. 3, p. 315-325, 1999.
- BECKER, G. A theory of marriage. *Journal of Political Economy*, v. 81, n.4, p. 813-846, 1973.
- BILLGER, S. *Does attending predominantly-female schools make a difference? Labor Outcomes for Women*, 2000. Mimeo.
- BLAU, F.; EHRENBERG, R. *Gender & family issues in the workplace*. Russel Sage Foundation, 1997.
- BRIEN, M.; LILLARD, L.; WAITE, L. Interrelated family-building behaviors: cohabitation, marriage and nonmarital conception. *Demography*, v. 36, n. 4, p. 531-551, 1999.
- BUMPASS, Larry L.; SWEET, James A. What is happening to the family? Interactions between demographics and institutional change. *Demography*, v. 27, n.4, p. 483-498, 1989.
- CARLSON, M.; DANZINGER, S. Cohabitation and the measurement of child poverty. *Review of Income and Wealth*, series 4, n. 2, p. 179-191, 1999.

- CASPER, L.; COHEN, P. How does POSSLQ measure up? Historical estimates of cohabitation. *Demography*, v. 37, n. 2, p. 237-245, 2000.
- GOLDIN, Claudia. *Understanding the gender gap: an economic history of american women*. Oxford University Press, 1990.
- HALLOCK, K.; HENDRICKS, W.; BROADBENT, E. Discrimination by gender and disability status: do worker perceptions match statistical measures? *Southern Economic Journal*, v. 65, n. 2, p. 245-263, 1998.
- HARKNESS, S.; WALDFOGEL, J. The family gap in pay: evidence from seven industrialized countries. CASE paper 29. *London School of Economics Working Paper*, 1999.
- HECKMAN, J. Sample selection bias as a specification error. *Econometrica*, v. 47, n.1, p. 153-162, 1979.
- HILL, M. The wage effects of marital status and children. *The Journal of Human Resources*, v. 14, n. 4, p. 579-594, 1979.
- HOFFMAN, R.; KASSOUF, A. L. Deriving conditional and unconditional marginal effects in log earnings equations estimated by Heckman's Procedure. *Applied Economics*, v. 37, p. 1303-1311, 2005.
- HU, W. Child support, welfare dependency, and women's labor supply. *Journal of Human Resources*, v. 34, n. 1, p. 71-103, 1999.
- JOSHI, H.; PACI, P.; WALDFOGEL, J. The wages of motherhood: better or worse? *Cambridge Journal of Economics*, v. 23, p. 543-564, 1999.
- KABEBERI-MACHARIA, J.; NYAMU, C. Marriage by affidavit: developing alternative laws on cohabitation in Kenya. In: EEKELAAR, J.; NHLAPO, T. *The changing family*. Oxford, UK: Hart Publishing, 1998.
- KORENMAN, S.; NEUMARK, D. Does marriage really make men more productive? *The Journal of Human Resources*, v. 24, n. 2, p. 282-307, 1991.
- _____. Marriage, motherhood, and wages. *The Journal of Human Resources*, v. 27, n. 2, p. 233-255, 1992.
- KUHN, P. Sex discrimination in labor markets: the role of statistical evidence. *The American Economic Review*, v. 77, n. 4, p. 567-583, 1987.
- LEE, Lung-Fei. Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables. *International Economic Review*, v. 19, n. 2, p. 415-433, 1978.
- LILLARD, L.; BRIEN, M.; WAITE, L. Premarital cohabitation and subsequent marital dissolution: a matter of self-selection. *Demography*, v. 32, n. 3, p. 437-457, 1995.
- MADDALA, G. S. Disequilibrium, self-selection, and switching models. In: *Griliches and Intriligator Handbook of Econometrics*, v. 3, p. 1633-1688, 1986.

- MOORE, W.; WILSON, R. The influence of children on the wage rates of married women. *Eastern Economic Journal*, v. 3, n. 3, 1982.
- NEAL, D. The economics of family structure. *National Bureau of Economic Research, Working Paper 8519*, 2001.
- OAXACA, R. Male-female wage differentials in urban labor markets. *International Economic Review*, v. 14, n. 3, p. 693-709, 1973.
- SAHA, A.; CAPPS Jr, O.; BYRNE, P. J. Calculating marginal effects in models for zero expenditures in household budgets using a Heckman-type Correction. *Applied Economics*, v. 29, 1311-1316, 1997.
- U. S. CENSUS BUREAU. *Statistical Abstract of the United States*. 1980 to 1999.
- VELLA, Francis. Estimating models with sample selection bias: a survey. *The Journal of Human Resources*, p. 127-169, Winter 1998.
- WAITE, L. Does marriage matter? *Demography*, v. 32, n. 4, p. 483-507, 1995.
- WALDFOGEL, Jane. Working mothers then and now: a cross-cohort analysis of the effects of maternity leave on women's pay. In: BLAU; EHRENBERG. *Gender & Family Issues in the Workplace*. Russell Sage Foundation, 1997.
- WHITE, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test of heteroskedasticity. *Econometrica*, v. 48, p. 817-838, 1980.
- WILLIS, R.; MICHAEL, R. Innovation in family formation: evidence on cohabitation in the United States. In: ERMISCH, J.; OGAWA, N. *The family, the market and the state in ageing societies*. Oxford: Clarendon Press, 1994.