Unemployment, Splitting Up and Spousal Income Replacement

by William Nilsson[1]

Department of Economics, Umeå University

In this study, the consequences of unemployment for a Swedish sample of couples are analyzed. The purpose is to estimate the possible income replacement that a spouse can provide. Unemployment can also affect the probability that the couples split up. Since not all couples remain in the analysis, a potential selection problem can occur. To deal with this problem, and also to take care of unobserved heterogeneity, a sample selection model for panel data is estimated. The results indicate that it is necessary to take into account the selection problem. A period in unemployment is found to be correlated with a higher female income only in the case of men who earned a fairly high income before becoming unemployed. Women who earned a fairly low income and were subject to a long period of unemployment are found to be compensated by a higher male income.

JEL: J12, J22, C33

Keywords: unemployment, divorce, spousal response, sample selection and panel data

1. Introduction

Going from employment to unemployment usually means a tighter personal budget. A single person who cannot find a job is likely to reduce some of his/her spending to compensate for the lower income. Individuals who live with a partner may have another option. Instead of just reducing consumption, it might be possible for the spouse to try to compensate the partner's loss in income by working more hours or finding a higher paid job. Having two sources of income is a way of not putting all one's eggs in the same basket. Further, compensating the partner's economic difficulties would be a way to handle risk within families by helping to reduce the family's vulnerability. Compensating behavior has, in the economic literature, been motivated by the idea that unemployment could provide the family with additional information about the labor market (Dynarski & Sheffrin, 1987 and Stephens, 2002). If the period in unemployment is expected to be long, or to be followed by more spells of unemployment, it is reasonable to expect that households will consider the possibilities for the spouse to try to compensate for the lower income.

Another branch of economic theory has focused on a different response arising from the new information accompanying unemployment. The literature on marital break ups predicts a higher risk of couples splitting when they are hit by economic shocks such as unemployment. (Becker *et al.*, 1977 and Charles & Stephens, 2004). Unemployment may, accordingly, affect both labor supply and the probability for the break up of the relationship. If the interest is to investigate the possible compensatory behavior on the part of a spouse, there could be a practical problem for empirical analysis if these two aspects are connected in some way. What if it is only couples where there is no, or only a small, need for an adjustment, who actually remain as couples? If unobservables affecting the decision to break up are correlated with unobservables influencing the

level of income, there could be a selection problem. Not taking into account that some couples will split up, could lead to biased estimates for the parameters of interest in explaining income.

The purpose of this study is to empirically investigate the possible economic response for the spouse arising from having an unemployed partner, while taking into account that some couples actually split up. In particular, the focus is on whether the spouse is able to reduce the vulnerability that can follow from unemployment. Earlier literature on labor supply responses for the spouse caused by the partner's unemployment, have underlined the importance of taking into account unobserved heterogeneity. (Maloney, 1991 and Bingley & Walker, 2001). For this reason, this study, uses an econometric model which deals with both the selection problem and unobserved heterogeneity. (Rochina-Barrachina, 1999)

The contributions of this study are the following: While earlier studies, such as Stephens (2002) and Charles & Stephens (2004), have pointed to the possible problem of marital break ups in empirical studies of economic responses on the part of the spouse arising from the partner's unemployment, this study models the probability that the couples will stay together. A sample selection model for panel data is used to address both the possible selection problem, and the unobserved heterogeneity, both of which can produce biased estimates. These kind of econometric models are still quite rare in empirical applications. This study investigates both a possible female response to male unemployment and vice versa. The latter has, to my knowledge, not been done in previous literature. The reason is that the female spouse has been seen as the second income earner. Limiting the empirical analysis in this way could, however, hide important results.

The paper is organized as follows. Section 2 presents the theoretical framework and gives examples of empirical studies. The empirical model is explained in section 3. The data is described in section 4 and the results from the empirical investigation are presented in section 5. Finally, concluding remarks can be found in section 6.

2. Theoretical framework and empirical studies

This theoretical overview discusses and brings together two different subfields in economic theory. First, the added worker effect, where additional information based on a spell of unemployment is expected to induce a labor supply response on the part of the spouse. The concept of the added worker effect usually refers to a situation with an unemployed individual, and a spouse who enters the labor market. In the Swedish case, it is typical that both spouses are working. Accordingly, investigating whether the individual's unemployment causes the spouse to enter the labor market is less relevant. The theory of the added worker effect is, however, still relevant in motivating a labor supply response. The second subfield in economic theory is the divorce literature where the additional information from a spell of unemployment is expected to increase the risk of divorce.

2.1 The added worker effect

The added worker effect refers to a situation where the individual's unemployment leads to an increase in the probability that the spouse will enter the labor force, or extend his/her hours of work. Two early studies of the subject are Mincer (1962) and Bowen & Finegan (1969).[2] In a life-

---

[2]Mincer (1962) finds an added worker effect while Bowen & Finegan (1969) find a discouraged worker effect. A discouraged worker effect means that the labor supply is reduced as a consequence of an unemployed spouse. The reason is that the husband's

cycle model, with intertemporal separability and perfect credit markets, Heckman & MaCurdy (1980) argue that the effect would be small. The reason is that the need for a response would be low with a perfect credit market and with a temporary reduction in income, which would be rather small in comparison with the lifetime income. Dynarski & Sheffrin (1987) later show, in a similar model, that this response could be present in the case where the job loss also includes additional information for the family. It is, for example, possible that the new situation of unemployment is due to a changed labor market in which the husband is likely to face long term difficulties in finding a new job. Stephens (2002) also focuses on the importance of new information in a household life-cycle model. In particular, he underlines the timing of information as the start of the process of the extended labor supply. It is possible that the response could start before the actual job loss, as information about, for example, downsizing or a plant closure, is likely to be available before the start of the unemployment spell. If the family believes that unemployment is very likely, the husband/wife could make his/her response very early, thus the effect observed at the time of the actual job loss could be rather small.

Many empirical studies of the added worker effect have highlighted the importance of taking into account unobserved heterogeneity as this may obscure the behavioral response. (See, for example, Maloney, 1991 and Bingley & Walker, 2001.) The idea is that a non random matching, i.e. assortative mating, could result in couples with similar skills and human capital. A highly educated woman may marry a highly educated man, and both would have quite good opportunities in the labor market. If assortative mating is important, observing an unemployed man could also mean a higher probability of observing a non-participating or unemployed

---

unemployment suggests that there is a difficult labor market. If there are costs associated with job search, it could be optimal to reduce searching if the probability of finding a job is low and, thus, the labor supply would also be reduced.

wife. Another possible reason that works against observing a response is that a bad local labor market could make employment difficult for both spouses. If, for example, an important local plant has to close, this could affect labor market opportunities for both spouses.

Maloney (1991) finds no evidence of an added worker effect. A husband's unemployment is not found to have an effect on either the wife's labor market participation or her employment. The study concludes by pointing to the importance of taking into account unobserved heterogeneity. Prieto-Rodríguez & Rodríguez-Gutiérrez (2003) use panel data in a study consisting of 11 European countries. They study the added worker effect in terms of labor market participation alternatively non-participation, i.e. they use a binary variable. The added worker effect is found to be unusual for the included countries. Bingley & Walker (2001) investigate the added worker effect not only by looking at participation or non-participation. They distinguish between those who work part time, full time and those who are not working at all. They also control for the possibility that non-participation is involuntary in that the woman actually wants to work but is unemployed. Bingley & Walker (2001) conclude that only looking at participation/non-participation may be inappropriate as the net effect could be either negative or positive. They also point to the importance of the duration of unemployment, as the effect was found to be different for women married to men who were unemployed for more than six month.

Stephens (2002) investigates the effect of displacement on the wife's hours of work. He allows for effects before the unemployment actually occurs, as information about a coming job loss can be available earlier. The effects on the wife's hours of work are also allowed for a number of years after the man's displacement, in order to observe the dynamics of the response. Another difference compared with earlier literature is that Stephens (2002) distinguishes between wives of low-income and high-income men who

become unemployed. The reason is that a low-income household could differ in response compared with a high-income household. The low-income household could be more likely to respond since they may have more difficulty to smooth expenditures when they experience earning losses. They might, on the other hand, be more used to a difficult labor market which should make a response less likely. Further, the family could already have taken the difficult situation into account and tried to increase its labor supply whenever possible.

The system for unemployment insurance can also lead to different responses, as this can affect the difference between the income before unemployment and during unemployment. A high-income family could, for example, be compensated to a lesser degree if the system has a maximum amount that can be received. The results from Stephens (2002) show no response to the husband's unemployment on the part of low-income households, while the wives of high-income men increase their hours of work. The results are in line with the hypothesis that high-income families tend to have larger income losses and that these necessitate a reaction on the part of the family. Further, the husband's loss of his job is more likely to have come unexpectedly for this group. In general, the response before displacement is estimated to be small, while the reaction is greater once job loss occurs. (Stephens, 2002)

Most of the discussion so far has been concerned with reasons for reacting to unemployment. Even though some sort of response to job loss would seem to be necessary it is, however, not obvious there will be a reaction. In the literature on the added worker effect, it is common to mention a discouraged worker effect. This refers to a situation of reduced labor supply as consequence of higher unemployment on the labor market. If search costs are high, it could be optimal to reduce job search because of the negative signal sent out by the partner's unemployment, i.e. that the

labor market situation is difficult. Further reasons for not observing a response are that it could be difficult to extend the hours of work if the person is, for example, already working full time. It could also be difficult to find additional or alternative ways to increase the income.

2.2 Unemployment and divorces

An important, and often cited, theoretical model concerning marital instability is developed in Becker *et al*. (1977). The main idea is that both spouses have expectations about future income earnings for themselves and their spouse. At the start of a marriage, couples often have different match quality with respect to their future stability concerning the risk that the marriage will break down. The match quality refers to, for example, similar experiences and common goals. Individuals update their expectations over time and evaluate the initial decision. Accordingly, new information which fails to match former expectations can cause a split. Note that both unanticipated negative and positive shocks can cause a divorce, since outside options are included in the decision.

In empirical investigations, the initial match quality is unobserved which complicates the procedure. It is appropriate to take into account unobserved heterogeneity since a split due to a initial bad match can be confused with, for example, unemployment of a spouse. Another difficulty, in the case of a correlation between divorce and the spouse's unemployment, is to know the direction of causality. Problems within a marriage could, for example, cause lower productivity and later unemployment. It is also possible that the correlation is not causally related, but rather connected to a third effect.

Weiss & Willis (1997) investigate how economic surprises are connected to marital splits. They estimate the surprise as the difference between

predicted earnings, based on earnings regressions from previous years, and actual observed earnings. A positive shock for the male reduces the probability of a divorce, while the opposite is found for females. Unemployment was not used in the study as a way of catching an unanticipated change that could cause a marital split. Charles & Stephens (2004) do, however, investigate the connection between job loss and divorce. They also study the effect of physical disability as another measure of a shock. In particular, they study the *first shock* that the married couples experience as it is in this situation that the most new information would appear. They do not find that disability experienced by either the male or the female partner has any effect on the probability of divorce. Job loss, experienced by the man or the woman, did, however, raise the probability of divorce significantly. Charles & Stephens (2004) also distinguish between job loss due to a plant closure and due to a layoff. Divorce is found to be more likely in the case of a layoff, which would suggest that there is more negative information for the spouse in a case where the husband or wife is laid off.

Another study concerning unemployment and the separation of couples is Kraft (2001). Special attention is paid to whether unemployment precedes the separation or not. This is the reason why he studies separations instead of divorces. Time to think for the couples, as well as the time involved in administrating a divorce usually means that the process takes some time until it is completed. The study includes the number of months in unemployment rather than a binary variable indicating the experience of unemployment during a year. A short period of unemployment, especially in cases of search unemployment, should not be assumed to be a severe shock that puts the relationship at risk. Further, in the study, lagged unemployment is used in an attempt to observe a possible causal relationship.

Kraft (2001) estimates a random effects probit model using a panel data set and, later, also a fixed effects logit model based on first differences. The results indicate that unemployment the previous year, measured in months, significantly increases the risk of separation.

2.3 Unemployment, divorce and spousal response

The theoretical base for the added worker effect and the divorce risk, point to the importance of new information as the reason for an increased labor supply, or an increased probability of divorce. In spite of this, it is nevertheless common to analyze the added worker effect using data on couples that remain together. Prieto-Rodríguez & Rodríguez-Gutiérrez (2003) use, for example, balanced panels of married couples. Stephens (2002), however, points to the potential problem with this approach and uses an unbalanced panel, in order to avoid basing the results only on families with long histories together. Charles & Stephens (2004) also use the potential problem involved in only estimating consumption or labor supply responses for families that remain together to motivate their study of job displacement and divorces. They argue that those couples that actually divorce are likely to have been hit by a relatively more severe drop in utility due to the shock. Those who stay together could be the couples that did not have to adjust that much. As stated in the introduction, this present study estimates both the decision to stay together and the potential income response for the spouse. The econometric model is described in the following section.

3. Empirical model

In the theoretical section, two different behavioral responses due to additional information in the form of a spouse's unemployment were discussed. Firstly, the added worker effect was explained, where unemployment could induce a compensating behavior on the part of the spouse. Secondly, the literature concerning divorces suggests that there is an increasing risk of a couple splitting up as a consequence of unemployment. If these responses are unrelated, estimating the added worker effect for the subsample of couples that remain together would not be associated with a selection problem. If, on the other hand, the behavioral responses are connected, neglecting the split would be likely to produce biased estimates. A sample selection problem could occur in this situation if unobservables that affect the decision to separate are correlated with unobservables that affect the income of the spouse. It is, for example, possible that couples, where considerable compensatory behavior is required on the part of the spouse, are more likely to split up. It is, therefore, potentially important to model both behavioral responses simultaneously.

3.1 Sample selection model for panel data

Taking into account these considerations, a sample selection model is estimated for the subsample that remains together. The response to an unemployed partner is assumed to be seen in terms of a change in the income of the spouse. Whether the response in income is due to more hours worked, or through obtaining a better paid job, is not investigated here. The reason for focusing on income, instead of on the number of hours worked, is that an income response directly indicates whether the spouse reduces the possible vulnerability resulting from a reduced income

caused by unemployment. These estimations take into account the potential selection problem that could occur if the calculations are only based on the subsample that stays together. The model is characterized as follows,

$$d^*_{it+1} = \boldsymbol{\gamma}' \mathbf{z}_{it} - \eta_i - u_{it}; \quad d_{it+1} = 1[d^*_{it+1} \geq 0] \tag{1}$$

$$y_{it} = \boldsymbol{\beta}' \mathbf{x}_{it} + \boldsymbol{\theta}' \mathbf{W}_{it} + \alpha_i + \varepsilon_{it} \quad if \ d_{it+1} = 1 \tag{2}$$

$i = 1,...,N$ and $t = 1,...,T$

where $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are all vectors of parameters that will be estimated. α_i, and η_i are unobserved time-invariant individual specific effects. These are allowed to be correlated with the explanatory variables, $\mathbf{z}_{it}$, $\mathbf{x}_{it}$ and $\mathbf{W}_{it}$. The explanatory variables in the selection equation (1) and the main equation (2) are allowed to have common elements. In fact, $\mathbf{W}_{it}$, indicating the spouse's unemployment, which is of main interest in the study, is also included in $\mathbf{z}_{it}$. ε_{it} and u_{it} are error terms. d_{it+1} is a binary variable that takes the value one if the couple is observed to remain together in year $t + 1$. The reason for investigating whether the couples stay together in $t + 1$ instead of t is to avoid confusing the direction of the effects. If unemployment and a split occur in the same period it could be problems within the relationship that proceed, and cause, unemployment. d^*_{it+1} is a latent propensity for staying together. In this application, y_{it} is labor income and work related social incomes. The estimation of these equations is complicated due to individual specific effects in both the selection equations and the main equations. Apart from that, an endogenous selection has to be taken into account. For these reasons, this study uses the sample selection model developed in Rochina-Barrachina

(1999) and later extended in Dustmann & Rochina-Barrachina (2000).[3]
The idea in Rochina-Barrachina (1999) is to deal with the unobserved
heterogeneity by estimating the equation that is of interest in terms of
differences. The selection problem is dealt with by including correction
terms constructed from bivariate probit models for different years.[4] The
sample selection model presented in Rochina-Barrachina (1999) relies on
three key assumptions.[5] Firstly, the regression function of η_i on z_i is
assumed to be linear. The individual effects are specified as a linear
projection of lags and leads of the explanatory variables in the selection
equation; $\eta_i = \delta_1' \mathbf{z}_{i1} + \ldots + \delta_T' \mathbf{z}_{iT} + c_i$, where c_i is a random effect.
Secondly, the errors in the selection equation, $v_{it} = u_{it} + c_i$, are assumed
to be normal $(0, \sigma_t^2)$. Finally, the errors $[(\varepsilon_{it} - \varepsilon_{is}), v_{it}, v_{is}]$ are trivariate

---

[3]Another option would have been the sample selection model developed in Kyriazidou
(1997) that has been used in Melenberg & van Soest (2001) and also Askildsen *et al.*
(2003). The model relies on that both the sample selection and the heterogeneity can be
removed by first difference for individuals with $d_{it} = d_{is} = 1$ and $\gamma' \mathbf{z}_{it} = \gamma' \mathbf{z}_{is}$. With
continuous variables, the latter equality is unlikely and the method uses instead kernel
weights to count observations, where the distance of $|\gamma' \mathbf{z}_{it} - \gamma' \mathbf{z}_{is}|$ is small, to a larger
extent. As noted in Dustmann & Rochina-Barrachina (2000), a practical problem could,
however, occur in applications. It is possible that observations where $\gamma' \mathbf{z}_{it}$ is close to
$\gamma' \mathbf{z}_{is}$ also have $\mathbf{x}_{it}$ close to $\mathbf{x}_{is}$. Observations that have small variation over time would,
accordingly, be weighted higher in the main equation. Both theories described in this
study underline the importance of unemployment as consisting of additional information.

Since unemployment for the spouse is an important part in both $\mathbf{z}_{it}$ and $\mathbf{x}_{it}$, the method
in Kyriazidou (1997) does not feel appropriate. Weighting the results on observations
which have small changes in the variables for unemployment could give misleading
estimates.

[4]If equation (2) were to be estimated in first difference for the subsample that are together
in both periods, t and s, β and θ would be consistently estimated under the condition that
$E(\varepsilon_{it} - \varepsilon_{is} | x_i, z_i, d_{it} = d_{is} = 1) = 0$, $s \neq t$, where
$x_i \equiv (x_{i1}, \ldots, x_{iT}, W_{i1}, \ldots, W_{iT})$ and $z_i \equiv (z_{i1}, \ldots, z_{iT})$. The idea of the
estimator, explained in Rochina-Barrachina (1999), is to parameterize the conditional
expectation, since the condition would not be fulfilled if a selection problem is present.

[5]Rochina-Barrachina (1999) also presents a less parametric estimator with semiparametric
individual effects. In this case, the first assumption and the estimation procedure are
different from the one described here.

normally distributed conditional on z_i, $\mathbf{x}_{it}$ and $\mathbf{W}_{it}$. With these assumptions and based on Rochina-Barrachina (1999), the regression equation for (3) is estimated in terms of differences,

$$y_{it} - y_{is} = \boldsymbol{\beta}'(\mathbf{x}_{it} - \mathbf{x}_{is}) + \boldsymbol{\theta}'(\mathbf{W}_{it} - \mathbf{W}_{is}) + l_{ts}\lambda_{its} + l_{st}\lambda_{ist} + e_{its} \tag{3}$$

The terms, λ_{its} and λ_{ist}, correct for the sample selection.[6] As in the method developed by Heckman (1979), the correction is based on estimates from a probit model for the selection equation. In this case, the correction terms are, however, estimated from a bivariate probit model for selection equation (1) in periods t and s. The correction terms from the first step are constructed as follows,

$$\lambda_{its} = \frac{\phi[\boldsymbol{\gamma}'_t\mathbf{z}_i]\cdot\Phi\left[\frac{[\boldsymbol{\gamma}'_s\mathbf{z}_i-\rho_{ts}\cdot\boldsymbol{\gamma}'_t\mathbf{z}_i]}{(1-\rho_{ts}^2)^{1/2}}\right]}{\Phi_2[\boldsymbol{\gamma}'_t\mathbf{z}_i,\boldsymbol{\gamma}'_s\mathbf{z}_i,\rho_{ts}]} \tag{4}$$

$$\lambda_{ist} = \frac{\phi[\boldsymbol{\gamma}'_s\mathbf{z}_i]\cdot\Phi\left[\frac{[\boldsymbol{\gamma}'_t\mathbf{z}_i-\rho_{ts}\cdot\boldsymbol{\gamma}'_s\mathbf{z}_i]}{(1-\rho_{ts}^2)^{1/2}}\right]}{\Phi_2[\boldsymbol{\gamma}'_t\mathbf{z}_i,\boldsymbol{\gamma}'_s\mathbf{z}_i,\rho_{ts}]} \tag{5}$$

where $\phi[\cdot]$ is the standard normal density function. $\Phi[\cdot]$ is the standardized univariate normal cumulative distribution function. Φ_2 refers to the bivariate normal cdf. Finally, ρ_{ts} is the estimated correlation coefficient between the error terms for the selection equations for t and s.

---

[6]The reason for that the correction terms are two, is that the differencing is made conditional on that the couples stay together in $t + 1$ and $s + 1$, i.e. two different years. See note 7 for further details.

The selection problem is dealt with by including the correction terms, and the individual specific heterogeneity is removed by taking differences. An ordinary least square regression for the subsample with $d_{it+1} = d_{is+1} = 1$ can, accordingly, yield consistent estimates for β and θ. Note that all possible combinations of years can be used for the estimation in differences. Further, note that the construction of the correction terms is not very different from the method developed by Heckman (1979), and the interpretation is made in the same way.[7] A Wald test of whether l_{ts} and l_{st} are jointly different from zero reveals if a selection problem is present. A positive sign, would, in this context, indicate that unobservables affecting the propensity to remain as a couple positively also affect the income positively. A negative sign, would suggest that unobservables affecting the propensity to remain as a couple positively, would affect the income negatively.

To obtain consistent estimates, it is, however, necessary that the assumption that all explanatory variables in both the selection equation and the main equation are strictly exogenous is fulfilled. The possibility that the time the spouse spends in unemployment could be affected by how easily his/her partner is able to adjust to the new situation, has been mentioned in the literature on the added worker effect. (Stephens, 2002). If he/she can easily find additional sources of income, the unemployed spouse can afford a longer period in unemployment and thus increase

---

[7] The lambda for Heckman's model, with only one period, would be $\phi[\gamma'\mathbf{z}_i]/\Phi[\gamma'\mathbf{z}_i]$, where $\phi[.]$ is the univariate standard normal density function and $\Phi[.]$ is the univariate cumulative distribution function. The first step in the construction of the lambda terms in Rochina-Barrachina's model is $\phi_2[\gamma'_t\mathbf{z}_i, \gamma'_s\mathbf{z}_i, \rho_{ts}]/\Phi_2[\gamma'_t\mathbf{z}_i, \gamma'_s\mathbf{z}_i, \rho_{ts}]$, where $\phi_2[.]$ is the bivariate normal density function. λ_{its} follows as a result of conditioning on $\gamma'_t\mathbf{z}_i$ being fixed and integrating over $\gamma'_s\mathbf{z}_i$. To get λ_{ist} the reverse is done. (Rochina-Barrachina, 1999).

his/her chances to find a better matched job with respect to his/hers human capital. If this were to be the case, the variables measuring the spouse's unemployment would not be strictly exogenous and the estimator would not be consistent.

In Dustmann & Rochina-Barrachina (2000), the estimator developed in Rochina-Barrachina (1999) is extended to allow for explanatory variables in the main equation that are not strictly exogenous. The parameters in the main equation can be estimated as,

$$b_{GMM} = \left\{ \sum_i \mathbf{x}_{its}^* \mathbf{q}_{its}' \ \hat{\Omega}^{-1} \ \sum_i \mathbf{q}_{its} \mathbf{x}_{its}^{*'} \right\}^{-1} \sum_i \mathbf{x}_{its}^* \mathbf{q}_{its}' \ \hat{\Omega}^{-1} \ \sum_i \mathbf{q}_{its} (y_{1it} - y_{1is}) \qquad (6)$$

where $\mathbf{x}_{its}^* \equiv [(x_{it} - x_{is}), (W_{it} - W_{is}), \lambda_{its}, \lambda_{ist}]'$ and $\mathbf{q}_{its} \equiv (q_i', \lambda_{its}, \lambda_{ist})$. $\mathbf{q}_{its}$ is a vector of instruments which includes the correction terms from the bivariate probit estimated for t and s. $\hat{\Omega}$ is first set to $\sum \mathbf{q}_{its} \mathbf{q}_{its}'$ and the estimator is, accordingly, an instrumental variables estimator.[8] The estimates are then used for the GMM estimator where $\hat{\Omega} = \sum \mathbf{q}_{its} \mathbf{q}_{its}' r_{its}^2$ and r_{its} are the residuals from the first step,

$$r_{its} = (y_{1it} - y_{1is}) - (x_{it} - x_{is}) b^{IV} - (W_{it} - W_{is}) b^{IV} - [l_{ts}^{IV} \lambda_{its} + l_{st}^{IV} \lambda_{ist}].$$

---

[8]Note that the structure is the same as the traditional instrumental variables estimation. The explanatory variables are just included in terms of differences, except λ_{its} and λ_{ist}, which are included as in equations (4) and (5).

3.2 Sensitivity analysis

The sample selection model described above could be flawed in this particular application. A problem occurs if the income, y_{it}, is a variable that should be included in equation (1). Including the income in the first step creates a simultaneity problem when the equation that is of interest is an income equation. Doing that, would, accordingly, yield biased estimates for the correction terms. The divorce literature has usually focused on unexpected income shocks, rather than on the income itself. However, a theoretical motivation for including the income as an explanatory variable for the probability that a couple will stay together is when couples are matched with respect to their earning capacity outside the household. According to this theory, a high earning male (female) matched to a low earning female (male) would be a stable couple since the couple could specialize in the labor outside, respective inside, household. (Becker *et al.* 1977). This theoretical motivation does, however, seems to be rather old-fashioned and less relevant for the Swedish case.

Another reason for including the income as an explanatory variable is where a high income is an important resource that could help the family to adapt to a new economic situation. This is also one reason, as will be explained in the next section, to construct the unemployment variables as interacted with the previous income. Despite this, it is still possible that the income could have an explanatory power in predicting which couples will stay together. Becker *et al.* (1977) find, for example, that divorces are more common at the tails of the earnings distribution. That is, low respective high earnings increase the risk for divorce. The suggested reason for this empirical result is that the earnings are more volatile and, accordingly, economic shocks are more common, at the tails of the distribution.

It is a complicated problem if income should be included in equation (1) and the dependent variable in equation (2) also is income. To my knowledge, there are not yet econometrical models that can handle this. A sensitivity analysis can, however, be performed in order to get an idea of whether the results, when the income is excluded, are flawed. An approximation of the correction terms is to directly include variables from the first step in the income equation instead of modelling the first step. In addition, interaction variables can be constructed to capture further heterogeneity. The conditional mean is, in this way, modelled directly, instead of through the bivariate probits. The results from both methods are compared. Before turning to the results, the description of the data is presented in the following section.

4. Data

The data for this study comes from the longitudinal database, Louise, for the years 1994-99.[9] The database is based on several different registries of the total population in Sweden. The population investigated is individuals born 1965. Even though the individuals are still quite young most of them should have entered into the labor market. The reason for examining a young population is that the first years after entering into the labor market could be difficult, and a spouse could be important in sharing the labor market risk. At the same time, it is more common for young couples to split up, which could make the selection problem larger. For this study, the individuals are matched with their identified spouses in 1994, and data for all years for those persons are also included. Information about whether the couples separate is also included for each of the subsequent years. The spouse could either be a married partner or a common-law spouse with common children.

---

[9]The database is described (in Swedish) in Bakgrundsfakta till Arbetsmarknads- och utbildningsstatistiken, 2002:2. Statistics Sweden.

Table 1. Descriptive statistics

Sample	Female		Male	
Variables	Mean	Std.err.	Mean	Std.err.
Staying together 1996	0.9317	0.2522	0.9236	0.2656
Staying together 1997	0.9022	0.2971	0.8912	0.3115
Staying together 1998	0.8746	0.3312	0.8598	0.3472
Staying together 1999	0.8483	0.3587	0.8339	0.3722
ln(labor income and work related social incomes)	11.7682	0.4405	12.2049	0.3764
Age difference, M1-2, (1=Male 1 or 2 years older)	0.2741	0.4461	0.3274	0.4693
Age difference, M3-5, (1=Male 3-5 years older)	0.2710	0.4445	0.2275	0.4192
Age difference, M≥6, (1=Male ≥6 years older)	0.2038	0.4028	0.0530	0.2240
Age difference, F1-2, (1=Female 1 or 2 years older)	0.0963	0.2951	0.1367	0.3435
Age difference, F≥3, (1=Female ≥3 years older)	0.0353	0.1845	0.1039	0.3052
Country (1=Both in couple born in same country)	0.9016	0.2979	0.9033	0.2956
Parents present in childhood (1=Both parents present at age 5, 10 and 15)	0.6878	0.4634	0.6961	0.4599
Region (1=Stockholm, Malmö or Gothenburg)	0.3425	0.4746	0.3410	0.4740
Months unemployed, 31-180, Q1. (Months unemployed, 31-180 days and income in 1st (lowest) quintile last year)	0.1812	0.8254	0.0606	0.4886
Months unemployed, >180, Q1. (Months unemployed, >180 days during year and income in 1st (lowest) quintile t-1)	0.6434	2.4933	0.4710	2.2180
Months unemployed, 31-180, Q2. (Months unemployed, 31-180 days and income in 2nd quintile last year)	0.1640	0.7811	0.0984	0.6201
Months unemployed, >180, Q2. (Months unemployed, >180 days during year and income in 2nd quintile t-1)	0.4017	1.9811	0.3977	2.0011
Months unemployed, 31-180, Q3. (Months unemployed, 31-180 days and income in 3rd quintile last year)	0.0408	0.3870	0.0804	0.5561
Months unemployed, >180, Q3. (Months unemployed, >180 days during year and income in 3rd quintile t-1)	0.0640	0.7990	0.1952	1.4010
Months unemployed, >30, Q4. (Months unemployed, >30 days during year and income in 4th quintile t-1)	0.0180	0.3691	0.1037	0.8627
Months unemployed, >30, Q5. (Months unemployed, >30 days during year and income in 5th (highest) quintile t-1)	0.0047	0.1962	0.0299	0.4601
Spouse, months unemployed, 31-180, Q1. (Months unemployed, 31-180 days and income in 1st (lowest) quintile t-1)	0.0702	0.5255	0.2416	0.9418
Spouse, months unemployed, >180, Q1. (Months unemployed, > 180 days during year and income in 1st (lowest) quintile t-1)	0.4999	2.2926	0.7978	2.7453
Spouse, months unemployed, 31-180, Q2. (Months unemployed, 31-180 days and income in 2nd quintile t-1)	0.0845	0.5810	0.1521	0.7452
Spouse, months unemployed, >180, Q2. (Months unemployed, > 180 days during year and income in 2nd quintile t-1)	0.3676	1.9149	0.3133	1.7386
Spouse, months unemployed, 31-180, Q3. (Months unemployed, 31-180 days and income in 3rd quintile t-1)	0.0610	0.4786	0.0318	0.3390
Spouse, months unemployed, >180, Q3. (Months unemployed, > 180 days during year and income in 3rd quintile t-1)	0.1588	1.2647	0.0533	0.7328
Spouse, months unemployed, >30, Q4. (Months unemployed, > 30 days during year and income in 4th quintile t-1)	0.0993	0.8633	0.0110	0.2816
Spouse, months unemployed, >30, Q5. (Months unemployed, > 30 days during year and income in 5th quintile t-1)	0.0246	0.4072	0.0041	0.1830

Table continues on the following page.

Table 1. Continues...

Sample	Female		Male	
Variables	Mean	Std.err.	Mean	Std.err.
Education, F2. (1=upper secondary school, less than 3 years)	0.6299	0.4828	0.6108	0.4876
Education F3. (1=post secondary school and post graduate education)	0.2621	0.4398	0.2492	0.4326
Education, F1, M2. (1=Female low education and male intermediate level of education)	0.1165	0.3209	0.0917	0.2886
Education, F1, M3. (1=Female low education and male high level of education)	0.0184	0.1343	0.0129	0.1127
Education, F2, M1. (1=Female intermediate education and male low level of education)	0.0597	0.2370	0.0724	0.2591
Education, F2, M2. (1=Female intermediate education and male intermediate level of education)	0.4037	0.4907	0.4355	0.4958
Education, F2, M3. (1=Female intermediate education and male high level of education)	0.1052	0.3068	0.1030	0.3039
Education, F3, M1. (1=Female high education and male low level of education)	0.0093	0.0958	0.0091	0.0950
Education, F3, M2. (1=Female high education and male intermediate level of education)	0.1096	0.3124	0.1073	0.3095
Education, F3, M3. (1=Female high education and male high level of education)	0.1385	0.3454	0.1328	0.3393
Child care, M all. (1=Male received all parental allowances)	0.0498	0.2176	0.0471	0.2119
Child care, F all. (1=Female received all parental allowances)	0.2437	0.4293	0.2230	0.4163
Child care, F 20. (1=Female received more than 0% and less than or equal to 20% of parental allowances)	0.0354	0.1847	0.0325	0.1774
Child care, F 40. (1=Female received more than 20% and less then or equal to 40% of parental allowances)	0.0578	0.2333	0.0522	0.2224
Child care, F 60. (1=Female received more than 40% and less then or equal to 60% of parental allowances)	0.0910	0.2876	0.0899	0.2860
Child care, F 80. (1=Female received more than 60% and less then or equal to 80% of parental allowances)	0.1365	0.3434	0.1417	0.3487
Child care, F100. (1=Female received more than 80% but less than 100% of child care allowance)	0.2652	0.4415	0.2908	0.4541
Not married (1=Couple not married)	0.3235	0.4678	0.3861	0.4869
Married, with children in household (1=Couple married and children present)	0.5942	0.4911	0.5138	0.4998
Number of children 0-3 years old	0.8587	0.6882	0.9404	0.6858
Number of children 4-6 years old	0.6029	0.6356	0.4896	0.6103
Number of children 7-10 years old	0.3396	0.5866	0.1760	0.4379
Number of children 11-17 years old	0.0506	0.2396	0.0352	0.2129
Spouses work in different municipalities (1=different municip.)	0.4355	0.4958	0.4445	0.4969
Study allowance (1=received study allowance > 5000 SEK)	0.0719	0.2583	0.0397	0.1953
Parental allowance (1=received parental allowance > 5000 SEK)	0.6073	0.4883	0.3700	0.4828

Note: Months in unemployment is measured as number of days/30. When 31-180 is indicated, the variables are coded to zero for individuals that were not unemployed or if unemployment was less than 31 days or more than 180 days. When >180 is indicated individuals with unemployment less than 181 are coded to zero. Note also that the variables are coded to zero unless the individual was in the mentioned quintile.

The sample for the analysis is divided into a male and a female sample, to see if the responses and effects are different. Self-employed individuals are excluded as they are able to control their incomes in a different way than employees. Further, only individuals and spouses who were alive until at least the end of 1999 are included. When observations with missing values are excluded, the female sample comprises 31 111 individuals and the male sample 23 620.[10]

Descriptive statistics can be found in *Table 1*. The year 1995 is used as an example for the variables that vary over time. For the two different samples, 93 percent and 92 percent of the couples in 1994 stayed together until 1996. In 1999, the last year of the panel, 85 percent and 83 percent were still together.

The income variable for the analysis is the sum of labor income and work related social incomes.[11] The latter type of income is such incomes or subsidies that, in principle are only paid when the individual is not working. This could, for example, be unemployment benefits, study allowances, sickness benefits, and parental allowances. The reason for including these kinds of incomes is that they are important parts of the formal insurance scheme to handle individual risk within the welfare state.

In the original data, unemployment is measured in terms of the number of days during the year that the individual was registered as unemployed. To obtain unemployment insurance in Sweden, the individual must be at the disposal of the labor market and searching for suitable jobs. In the empirical analysis, variables are constructed depending, as described below, on the time in unemployment and the labor income the previous year. The reason for taking into account previous income is that in Sweden

---

[10]The reason for the smaller male sample is that fewer male individuals, that are born 1965, have an identified spouse in 1994.

[11]The incomes are deflated to the 2001 price level by using the consumer price index.

the amount of money an individual receives as unemployment insurance depends on former earned income. Up to a certain level, 80 percent of the income is paid by the government as insurance.[12] A maximum amount is set by the government where, above that level, no more income replacement is granted. This means that the income loss can be substantial for high income groups, whereas low income earners do not reach the ceiling, and are accordingly paid 80 percent of their former earnings.

The unemployment variables are constructed to take into account this possible nonlinear relationship in the response. The first step in constructing the unemployment variables is to order the spouses in quintiles with respect to their income the previous year.[13] Five dummy variables are constructed to indicate in which quintile the spouse is included. The second step is to create dummy variables depending on the length in unemployment. In the first dummy variable for the length, individuals unemployed between 31 and 180 days are coded as one and all other cases as zero. In the second dummy variable, individuals unemployed more than 180 days are coded as one and all other cases as zero. The third step is to measure the time in unemployment in months instead of days. The months spent in unemployment are calculated as the number of days divided by 30, and the move to months is done only to ease the interpretation of the estimates. With all these dummy variables, and the time in unemployment measured in months, the final step is to make interactions to receive the variables that are used in the estimations. The first variable is the dummy variable for the first (lowest) quintile

[12]1996 and 1997 the level of compensation was 75% of the previous income. The rest of the years between 1994 and 1999 the level of compensation was 80%. The maximum amount that could be received was 564 SEK/day between 1994 and 1997, and 580 SEK/day in 1998 and 1999. The compensation only refers to 5 days for each week of unemployment. (Bakgrundsfakta till Arbetsmarknads- och utbildningsstatistiken, 2002:2).
[13]In the first step, where the probability to stay together is estimated, variables for unemployment of the individual are also used. The manner used to construct the variables is the same.

multiplied by the dummy variable for unemployment between 31 and 180 days and the time in unemployment measured in months. The second variable is the dummy variable for the first (lowest) quintile multiplied with the dummy variable for unemployment longer than 180 days and the time in unemployment measured in months. The same procedure is made for all quintiles. Very few individuals in quintile four and five are observed as unemployed the following year. The two variables for the time in unemployment constructed for quintile four are, for this reason, added to make one variable. The final variable includes the unemployment measured in months, when unemployment was more than 30 days, and the spouse was in the fourth income quintile the previous year. The same is done for the fifth income quintile. The unemployment variables are now designed to allow for a different response for the spouse, depending on both the length in unemployment and the previous income.

Education dummies are used to capture whether different combinations of the level of education between the spouses are important for whether the couples stay together. Another set of dummies controls for how the childcare allowance is shared between the spouses. A woman who has the main responsibility for taking care of the children could, for example, be less independent when it comes to working outside the home. The number of children within different age groups is also included as several different variables. The reason is to capture that children could be seen as an investment in the relationship.

An alternative specification is also estimated to see whether the level of education for the individual is important for a possible income response arising from having an unemployed spouse. Education is interacted with the unemployment variables mentioned above.

A response could, accordingly, be different depending on the level of education. The possibility, but also the need to make a response could, for example, be different depending on the level of education.

The variables included as control variables in equation (2) are the number of children aged 0-3, 4-6, 7-10 and 11-17 years old and dummy variables indicating whether the individual lives in a big city, i.e. Stockholm, Malmö or Gothenburg. Two dummy variables are also included for different levels of education, where the reference case is compulsory schooling. Another dummy variable indicates whether the individual received a study allowance of more than 5000 SEK during the year. The idea is that this would indicate that the individual is engaged in higher education and, accordingly, would earn a lower income during the year. Another dummy variable, based on the same idea, indicating that the individual received a parental allowance of more than 5000 SEK during the year is also included.

5. Results

In the theoretical section a potential selection problem was discussed concerning the fact that not all couples stay together. An individual's unemployment can affect both his/here spouse's labor supply and the probability that the couple will break up. Section 3 described a sample selection model for panel data that also accounted for unobserved heterogeneity. The results from this model are presented here. The presentation starts with the selection step for both the female and the male sample. After that, a potential response for women arising from an unemployed partner is described. Before repeating this analysis for the male sample, a sensitivity analysis is performed.

The data used in the study are collected from the period 1994 to 1999. The construction of the variables for unemployment uses the quintile of the income the previous year. Therefore, the first year of the panel cannot be used in the regression. The year for the potential split up of the couple is also the following year, $t + 1$. The equation that is of interest can, accordingly, be estimated using information from 1995 to 1998. The correction terms for the sample selection are estimated using bivariate probits for different combination of years for whether the couple stays together $t + 1$. With bivariate probits for a total of six different combinations of years, the included variables produce 12 different estimates for the parameters. The results from the bivariate probits for both the female and the male samples are summarized[14] in *Table 2*.

Table 2 includes the number of cases in which the coefficients are found to be significantly different from zero at a 5 percent significance level in double sided tests. Of course, a parameter cannot be estimated significantly different from zero more than 12 times. The overview indicates that unemployment is connected negatively to staying together for the sample investigated. In particular, a low level of income the previous year combined with unemployment seems to reduce the probability that the couples will stay together. Having had both parents present in their own childhood increases the probability that the couples will remain together.[15] Higher education, in general, seems to increase the probability of remaining together compared with the reference case where both partners had a low level of education.

---

[14]The complete estimates for all bivariate probits can be requested from the author.
[15]The variable takes the value of one, if the same social parents were present 1970, 1975 and 1980, and zero otherwise. It is, accordingly, not necessary that the parents are the biological parents.

Table 2. Summarized information from bivariate probits, 1 = stay together t + 1

Sample	Female		Male	
	Number of coefficients significantly different from zero at the 5% significance level			
Variables, see *Table 1* for description.	positive	negative	positive	negative
Constant	5	3	4	5
Age difference, M1-2, (1=Male 1 or 2 years older)	1	0	0	0
Age difference, M3-5, (1=Male 3-5 years older)	0	0	0	5
Age difference, M≥6, (1=Male ≥6 years older)	0	2	0	9
Age difference, F1-2, (1=Female 1 or 2 years older)	0	7	0	1
Age difference, F≥3, (1=Female ≥3 years older)	0	12	0	1
Country (1=Both in couple born in same country)	12	0	12	0
Parents present in childhood	12	0	12	0
Region (1=Stockholm, Malmö or Gothenburg)	0	12	0	1
Months unemployed, 31-180 days, Quintile 1.	0	9	0	8
Months unemployed, >180, Q1.	0	9	0	12
Months unemployed, 31-180, Q2.	0	2	0	9
Months unemployed, >180, Q2.	0	11	0	12
Months unemployed, 31-180, Q3.	0	2	0	2
Months unemployed, >180, Q3.	0	0	0	9
Months unemployed, >30, Q4.	0	0	0	3
Months unemployed, >30, Q5.	0	2	0	2
Spouse, months unemployed, 31-180, Q1.	0	9	0	10
Spouse, months unemployed, >180, Q1.	0	12	0	10
Spouse, months unemployed, 31-180, Q2.	0	9	0	8
Spouse, months unemployed, >180, Q2.	0	12	0	4
Spouse, months unemployed, 31-180, Q3.	0	4	0	0
Spouse, months unemployed, >180, Q3.	0	9	0	0
Spouse, months unemployed, >30, Q4.	0	4	0	0
Spouse, months unemployed, >30, Q5.	0	6	1	0
Education, Female1 (low), Male2 (intermediate)	2	0	12	0
Education, Female1 (low), Male3 (high)	2	0	10	0
Education, F2, M1.	0	1	1	0
Education, F2, M2.	8	0	12	0
Education, F2, M3.	12	0	12	0
Education, F3, M1.	0	0	0	0
Education, F3, M2.	12	0	12	0
Education, F3, M3.	12	0	12	0
Child care, Male got all child allowance	0	11	0	12
Child care, Female got all child allowance	0	11	0	12
Child care, F 20.	0	7	0	12
Child care, F 40.	0	9	0	11
Child care, F 60.	0	8	0	12
Child care, F 80.	0	6	0	11
Child care, F100.	0	3	0	11

Table continues on the following page.

Table 2. Continues...

Sample	Female		Male	
	Number of coefficients significantly different from zero at the 5% significance level			
Variables, see *Table 1* for description.	positive	negative	positive	negative
Not married (1=Couple not married)	12	0	12	0
Married, with children in household	12	0	12	0
Number of children 0-3 years old	12	0	12	0
Number of children 4-6 years old	6	1	12	0
Number of children 7-10 years old	3	8	11	0
Number of children 11-17 years old	1	3	7	0
Spouses work in different municipalities	0	11	0	7

A set of dummy variables are used to see whether the manner in which child care transfers are shared between the couple influences the probability of staying together.[16] The reference case is couples that do not receive any parental allowances and, accordingly, do not have any children. The overall impression is that the way in which the transfers are shared is not important, but rather what matters is whether the couples receive a child care allowance at all. Receiving an allowance seems to reduce the probability that the couple will remain together.

If this selection step is repeated, and the male and female incomes are also included as explanatory variables, they are shown to have significant effects on the probability of staying together. For the female sample, the coefficient for the woman's income is significantly negative in eight out of 12 cases.[17] The coefficient for the man's income is significantly positive in

---

[16]When a baby is born, couples in Sweden have a total of 450 days of parental leave so that they can stay at home with the child and are compensated based on the previous earned income. These days can be shared between the mother and the father, but from 1995 one month is specifically assigned to each parent. Compensation is also available so that a parent can take care of a sick child. Further, having a child in Sweden also means a monthly child allowance is paid to the parents. Note that the dummy variables concern the share of the compensation and not the share of time.

[17]For the male sample, the coefficient for male income was significantly positive in 10 out of 12 cases. The female income was significantly positive in two cases and significantly negative in one case. The difference between the samples is most likely due to differences in the ages of the female spouses. It seems that the female income is more important in the sample where the female spouse is a few more years older. These results can be requested from the author.

nine out of 12 cases. This result is interesting, but it also means that the results in the next step have to be interpreted with care. The sensitivity analysis seems to be important.

Table 3. Estimates. Dependent variable, ln(income)

| Sample: Female   Estimation method: GMM | | | | Sensitivity |
Variables	Coefficients	Std.errors	t-value	analysis
Region (1=Stockholm, Malmö or Gothenburg)	-0.0029	0.0066	-0.44	
Spouse, months unemployed, 31-180 days, Quintile 1.	-0.0171	0.0179	-0.95	
Spouse, months unemployed, >180 days, Quintile 1.	0.0009	0.0027	0.34	
Spouse, months unemployed, 31-180 days, Quintile 2.	-0.0004	0.0169	-0.02	
Spouse, months unemployed, >180 days, Quintile 2.	-0.0021	0.0043	-0.48	
Spouse, months unemployed, 31-180 days, Quintile 3.	0.0308*	0.0179	1.72	
Spouse, months unemployed, >180 days, Quintile 3.	-0.0035	0.0063	-0.55	
Spouse, months unemployed, >30 days, Quintile 4.	0.0202***	0.0074	2.71	+
Spouse, months unemployed, >30 days, Quintile 5.	-0.0828***	0.0174	-4.77	
Number of children 0-3 years old	-0.1216***	0.0073	-16.63	---
Number of children 4-6 years old	-0.0948***	0.0080	-11.82	---
Number of children 7-10 years old	-0.1018***	0.0104	-9.79	---
Number of children 11-17 years old	-0.1038***	0.0137	-7.58	---
Education, F2. (1=upper secondary school, less than 3 years)	-0.0112	0.0350	-0.32	
Education, F3. (1=post secondary school and post graduate education)	-0.1307***	0.0366	-3.57	
Study allowance (1=study allowance > 5000 SEK	-0.2589***	0.0089	-29.22	
Parental benefits (1=parental benefits > 5000 SEK	-0.2307***	0.0082	-28.14	---
Time dummy 97 (1=1997)	0.0155***	0.0053	2.94	+++
Time dummy 98 (1=1998)	0.0823***	0.0072	11.39	+++
Time dummy 99 (1=1999)	0.1615***	0.0087	18.57	
$\lambda_{95,96}$	0.7831***	0.0742	10.55	
$\lambda_{96,95}$	-0.1773***	0.0317	-5.60	
$\lambda_{95,97}$	0.2369***	0.0578	4.10	
$\lambda_{97,95}$	-0.0068	0.0232	-0.30	
$\lambda_{95,98}$	0.2520***	0.0529	4.76	
$\lambda_{98,95}$	-0.0378**	0.0179	-2.12	
$\lambda_{96,97}$	0.1118***	0.0424	2.64	
$\lambda_{97,96}$	-0.1138***	0.0288	-3.95	
$\lambda_{96,98}$	0.0302	0.0325	0.93	
$\lambda_{98,96}$	-0.0681***	0.0201	-3.39	
$\lambda_{97,98}$	0.0151	0.0261	0.58	
$\lambda_{98,97}$	-0.0175	0.0202	-0.87	

Wald test whether λ-terms are joint significantly different from zero;   152.9644***

Note: Income is measured as sum of labor income and work related social incomes. See *Table 1* for further explanations of variables. Estimation is made in differences. Standard errors are corrected for first step estimations to construct λ-terms. Coefficients that are significantly different from zero at 10, 5 and 1%-level are marked with *, ** and ***. Coefficients that are positively (negatively) significantly different from zero at 10, 5 and 1%-level in the sensitivity regression are marked with + (-), ++ (--) and +++ (---). See text for explanation of the sensitivity analysis.

The bivariate probits are used to correct for the possible sample selection problem. The results for the income equation for the female sample are included in *Table 3*.

The estimates for each possible pair of differences for the years 1995-98 are combined with a minimum distance procedure.[18] A weighting matrix, based on the different estimated covariance matrixes, is used. (See Rochina-Barrachina, 1999 for a more detailed explanation on the minimum distance procedure.) The standard errors in *Table 3* are corrected for the variability induced from the estimated correction terms from the bivariate probit models. The GMM estimator is used as the unemployment variables are rejected as strictly exogenous with a Wu-test.[19] The principle for choosing the set of instruments is the lags and leads of the variables used in the selection step and the estimated correction terms for each pair of time periods. The unemployment variables are, however, not used as instruments for the period $t-1$, because of the risk that the variables could have an explanatory power in the income equation, implying inconsistent estimates. The individual's unemployment is, for the same reason, never used as an instrument. A Wald test for the joint significance of the correction terms shows that a selection problem is present.[20]

The results show that male unemployment, in general, does not influence the income of the female spouse. However, for those men included in the fourth quintile the previous year, unemployment seems to be positively correlated with the spouse's income. The coefficient is significantly

---

[18]The coefficients for the lambda terms are not combined with this method. The coefficients for the lambda terms are all included in the tables.

[19]The Wu-test is equivalent to the Hausman-test, although easier to implement (Greene, 2003). The Wald statistic is estimated to be 15.60 compared to the critical value of 14.07 for the 5% significance level with seven degrees of freedom. Accordingly, the assumption that the variables measuring unemployment are strictly exogenous is rejected.

[20]With four periods to use, a total of six different combinations of periods results in 12 correction terms. The Wald test statistic is estimated to be 153.96, which is larger than the critical value of 21.03 in the chi-square distribution with 12 degrees of freedom for the 5% significance level.

different from zero at the 1 percent significance level. For the highest income quintile the previous year, the coefficient is estimated to be negative and significant different from zero at the 1 percent significance level. Apart from these significant responses, only a short time in unemployment for individuals in the third quintile has a tendency to illustrate a compensatory behavior. The coefficient is significantly different from zero at the 10 percent significance level. The sensitivity analysis indicates that the possible response is quite limited.[21] A significant coefficient is repeated only for the fourth quintile. The level of significance is, however, only at the 10 percent level.

The parameters for the unemployment variables included are, in some cases, estimated with large standard errors. This is typical if the instruments are weakly correlated with the endogenous variables. It is, accordingly, possible that the estimates are inconsistent due to a correlation between the instruments and the residuals in the original equation. This can occur even if the correlation is small. (Bound, Jaeger & Baker, 1995).

Table 4 includes estimates from the sample selection model where the unemployment variables are interacted with the level of education. This is done to allow different responses depending on the educational level, and the possible need and opportunity to make a response. The interaction is only done for the first three quintiles as there are quite few individuals included in quintiles four and five of the income distribution the previous

---

[21]The sensitivity analysis is done by including the time-variant variables in the first step directly, instead of using the lambda terms. Interaction terms are also constructed from some of the time-invariant variables combined with time-variant variables. The interaction variables are; (Female older than male=1)*(Study allowance > 5000 SEK = 1), (Female older than male=1)*(Parental allowance > 5000 SEK = 1), (Female older than male=1)*(Female received all parental allowance=1), (Female received all parental allowance=1)*(Education level 1 for individual=1), (Female received all parental allowance=1)*(Education level 3 for individual=1). The results from the sensitivity estimations can be requested from the author.

year and later unemployed. The sample selection model is used with the GMM estimation procedure. The combination of a high level of education, together with long term unemployment for a male spouse previously belonging to the third quintile, indicates a negative response on the part of the female spouse. The sensitivity analysis does, however, not repeat this result. Unemployment for the fourth quintile seems, as before, to have a significant impact on the female income.

Table 4. Estimates. Dependent variable, ln(income).

Sample: Female Estimation method: GMM				Sensitivity
Variables	Coefficients	Std.errors	t-value	Analysis
Region (1=Stockholm, Malmö or Gothenburg)	-0.0019	0.0069	-0.28	
Educ., F1* Spouse, months unempl., 31-180 days, Q1.	0.0200	0.0476	0.42	
Educ., F1* Spouse, months unempl., >180 days, Q1.	0.0046	0.0056	0.82	
Educ., F1* Spouse, months unempl.., 31-180 days, Q2.	-0.0413	0.0353	-1.17	
Educ., F1* Spouse, months unempl., >180 days, Q2.	0.0063	0.0087	0.72	--
Educ., F1* Spouse, months unempl., 31-180 days, Q3.	-0.0236	0.0485	-0.49	
Educ., F1* Spouse, months unempl., >180 days, Q3.	0.0143	0.0132	1.08	
Educ., F2* Spouse, months unempl., 31-180 days, Q1.	-0.0290	0.0246	-1.18	
Educ., F2* Spouse, months unempl., >180 days, Q1.	-0.0013	0.0038	-0.34	
Educ., F2* Spouse, months unempl., 31-180 days, Q2.	0.0237	0.0237	1.00	
Educ., F2* Spouse, months unempl., >180 days, Q2.	-0.0057	0.0053	-1.06	
Educ., F2* Spouse, months unempl., 31-180 days, Q3.	0.0295	0.0231	1.28	
Educ., F2* Spouse, months unempl., >180 days, Q3.	0.0030	0.0077	0.39	
Educ., F3* Spouse, months unempl., 31-180 days, Q1.	0.0052	0.0317	0.17	
Educ., F3* Spouse, months unempl., >180 days, Q1.	0.0020	0.0064	0.31	
Educ., F3* Spouse, months unempl., 31-180 days, Q2.	0.0433	0.0398	1.09	
Educ., F3* Spouse, months unempl., >180 days, Q2.	0.0165	0.0114	1.44	++
Educ., F3* Spouse, months unempl., 31-180 days, Q3.	-0.0018	0.0348	-0.05	
Educ., F3* Spouse, months unempl., >180 days, Q3.	-0.0473***	0.0160	-2.96	
Spouse, months unemployed, >30 days, Quintile 4.	0.0182**	0.0078	2.35	++
Spouse, months unemployed, >30 days, Quintile 5.	-0.0700***	0.0173	-4.05	
Number of children 0-3 years old	-0.1269***	0.0076	-16.63	---
Number of children 4-6 years old	-0.0986***	0.0084	-11.76	---
Number of children 7-10 years old	-0.1026***	0.0109	-9.44	---
Number of children 11-17 years old	-0.1018***	0.0143	-7.11	---
Educ., F2. (1=upper secondary school, less than 3 years)	-0.0023	0.0423	-0.06	
Educ., F3. (1=post secondary school and post grad. educ.)	-0.1315***	0.0445	-2.96	
Study allowance (1=study allowance > 5000 SEK	-0.2532***	0.0092	-27.65	
Parental benefits (1=parental benefits > 5000 SEK	-0.2237***	0.0086	-25.95	---
Time dummy 97 (1=1997)	0.0165***	0.0058	2.86	+++
Time dummy 98 (1=1998)	0.0846***	0.0076	11.21	+++
Time dummy 99 (1=1999)	0.1632***	0.0091	17.94	+++

Table continues on the following page.

Table 4. Continues...

Sample: Female	Estimation method: GMM			Sensitivity
Variables	Coefficients	Std.errors	t-value	Analysis
$\lambda_{95,96}$	0.7279***	0.0824	8.84	
$\lambda_{96,95}$	-0.1662***	0.0362	-4.59	
$\lambda_{95,97}$	0.2318***	0.0631	3.67	
$\lambda_{97,95}$	-0.0191	0.0253	-0.76	
$\lambda_{95,98}$	0.2455***	0.0564	4.36	
$\lambda_{98,95}$	-0.0473**	0.0187	-2.52	
$\lambda_{96,97}$	0.1156**	0.0508	2.28	
$\lambda_{97,96}$	-0.1225***	0.0341	-3.60	
$\lambda_{96,98}$	0.0360	0.0422	0.85	
$\lambda_{98,96}$	-0.0688***	0.0243	-2.84	
$\lambda_{97,98}$	0.0159	0.0277	0.57	
$\lambda_{98,97}$	-0.0200	0.0211	-0.95	

Wald test whether λ-terms are joint significantly different from zero;   98.5467***

Note: See *Table 1* for explanations of variables. Coefficients that are significantly different from zero at 10, 5 and 1%-level are marked with *, ** and ***. Coefficients that are positively (negatively) significantly different from zero at 10, 5 and 1%-level in the sensitivity regression are marked with + (-), ++ (--) and +++ (---). See text for explanation of the sensitivity analysis.

To see whether female unemployment is connected with a male income response the male sample is used. The results are included in *Table 5*.

For the male sample the variables for female unemployment are rejected as strictly exogenous and again the GMM-version is used.[22] A small but significant effect, at the 5 percent level, is found for a long period of female unemployment when the female was in the second quintile of the earnings distribution the previous year. This is the only relationship concerning the spouse's unemployment that is confirmed by the sensitivity analysis.

---

[22]The Wald statistic is estimated to be 14.25 and accordingly the hypothesis that the unemployment variables are exogenous is rejected. (The critical value, for the 5% significance level, in the chi-square distribution with 7 degrees of freedom is 14.07.)

Table 5. Estimates. Dependent variable, ln(income).

Sample: Male Estimation method: GMM				Sensitivity
Variables	Coefficients	Std.errors	t-value	Analysis
Region (1=Stockholm, Malmö or Gothenburg)	-0.0116**	0.0053	-2.16	--
Spouse, months unemployed, 31-180 days, Q1.	-0.0101	0.0066	-1.52	
Spouse, months unemployed, >180 days, Q1.	-0.0055***	0.0015	-3.62	
Spouse, months unemployed, 31-180 days, Q2.	-0.0032	0.0093	-0.34	
Spouse, months unemployed, >180 days, Q2.	0.0083**	0.0035	2.38	++
Spouse, months unemployed, 31-180 days, Q3.	0.0166	0.0157	1.06	
Spouse, months unemployed, >180 days, Q3.	-0.0036	0.0077	-0.47	
Spouse, months unemployed, >30 days, Q4.	-0.0046	0.0135	-0.34	
Spouse, months unemployed, >30 days, Q5.	0.0481	0.0515	0.93	
Number of children 0-3 years old	0.0177***	0.0061	2.88	+
Number of children 4-6 years old	0.0057	0.0067	0.86	
Number of children 7-10 years old	-0.0133	0.0084	-1.58	
Number of children 11-17 years old	-0.0362***	0.0113	-3.20	-
Education, M2. (1=upper secondary school, less than 3 years)	0.0562	0.0442	1.27	
Education, M3. (1=post secondary school and post graduate education)	-0.0098	0.0466	-0.21	
Study allowance (1=study allowance > 5000 SEK	-0.3745***	0.0148	-25.27	---
Parental benefits (1=parental benefits > 5000 SEK	0.0019	0.0091	0.21	
Time dummy 97 (1=1997)	0.0657***	0.0032	20.34	+++
Time dummy 98 (1=1998)	0.1287***	0.0052	24.75	+++
Time dummy 99 (1=1999)	0.2023***	0.0068	29.80	+++
$\lambda_{95,96}$	-0.0373	0.0391	-0.95	
$\lambda_{96,95}$	-0.0714***	0.0202	-3.54	
$\lambda_{95,97}$	-0.1245***	0.0359	-3.47	
$\lambda_{97,95}$	-0.0360**	0.0172	-2.09	
$\lambda_{95,98}$	-0.1243***	0.0336	-3.70	
$\lambda_{98,95}$	-0.0462***	0.0167	-2.77	
$\lambda_{96,97}$	-0.0464**	0.0214	-2.17	
$\lambda_{97,96}$	0.0022	0.0179	0.13	
$\lambda_{96,98}$	-0.0443**	0.0214	-2.07	
$\lambda_{98,96}$	-0.0191	0.0170	-1.12	
$\lambda_{97,98}$	-0.0190	0.0171	-1.11	
$\lambda_{98,97}$	-0.0107	0.0161	-0.67	

Wald test whether λ-terms are joint significantly different from zero;   43.0889***

Note: Income is measured as sum of labor income and work related social incomes. See *Table 1* for further explanations of variables. Estimation is made in differences. Standard errors are corrected for first step estimations to construct λ-terms. Coefficients that are significantly different from zero at 10, 5 and 1%-level are marked with *, ** and ***. Coefficients that are positively (negatively) significantly different from zero at 10, 5 and 1%-level in the sensitivity regression are marked with + (-), ++ (--) and +++ (---). See text for explanation of the sensitivity analysis.

The difference between the male and female samples concerning the importance of children for the income is evident. For the female sample, children are more important in affecting the income. Even though there is one case in which children have a significantly negative effect on the male income, it is far less severe in magnitude. In fact, having children between 0 and 3 years old, is correlated with a higher male income.

Results for the male sample, when the unemployment variables are interacted with different levels of education, are included in *Table 6*. Even though few of the unemployment variables are significantly different from zero, there seems to be a difference depending on the level of education. A positive response is found in the case of short term unemployment for the first quintile when the male had a low education. The coefficient is significant at the 5 percent significance level. A positive response is also found for long term unemployment for the second quintile when the male had a low education. The coefficient is significant at the 1 percent significance level. Both these responses are confirmed by the sensitivity analysis. Unemployment for the first quintile is found to be negatively correlated with the male income, when the man is highly educated. The coefficients are significantly different from zero at the 1 percent significance level both for long term and short term unemployment. The sensitivity analysis confirms these estimates.

Table 6. Estimates. Dependent variable, ln(income).

Sample: Male Estimation method: GMM Variables	Coefficients	Std.errors	t-value	Sensitivity Analysis
Region (1=Stockholm, Malmö or Gothenburg)	-0.0102*	0.0057	-1.80	--
Educ., M1* Spouse, months unempl., 31-180 days, Q1.	0.0281**	0.0123	2.28	+
Educ., M1* Spouse, months unempl., >180 days, Q1.	-0.0042	0.0036	-1.18	
Educ., M1* Spouse, months unempl., 31-180 days, Q2.	0.0115	0.0217	0.53	
Educ., M1* Spouse, months unempl., >180 days, Q2.	0.0157***	0.0057	2.75	+
Educ., M1* Spouse, months unempl., 31-180 days, Q3.	-0.0078	0.0206	-0.38	
Educ., M1* Spouse, months unempl., >180 days, Q3.	0.0044	0.0097	0.46	
Educ., M2* Spouse, months unempl., 31-180 days, Q1.	-0.0015	0.0096	-0.16	
Educ., M2* Spouse, months unempl., >180 days, Q1.	0.0011	0.0019	0.55	
Educ., M2* Spouse, months unempl., 31-180 days, Q2.	0.0019	0.0111	0.17	
Educ., M2* Spouse, months unempl., >180 days, Q2.	0.0025	0.0041	0.62	
Educ., M2* Spouse, months unempl., 31-180 days, Q3.	0.0237	0.0216	1.10	
Educ., M2* Spouse, months unempl., >180 days, Q3.	-0.0058	0.0089	-0.64	
Educ., M3* Spouse, months unempl., 31-180 days, Q1.	-0.0668***	0.0145	-4.60	---
Educ., M3* Spouse, months unempl., >180 days, Q1.	-0.0212***	0.0055	-3.85	-
Educ., M3* Spouse, months unempl., 31-180 days, Q2.	-0.0256	0.0205	-1.25	
Educ., M3* Spouse, months unempl., >180 days, Q2.	-0.0112	0.0092	-1.22	++
Educ., M3* Spouse, months unempl., 31-180 days, Q3.	0.0539	0.0398	1.36	
Educ., M3* Spouse, months unempl., >180 days, Q3.	-0.0079	0.0164	-0.48	
Spouse, months unemployed, >30 days, Quintile 4.	-0.0287**	0.0142	-2.02	
Spouse, months unemployed, >30 days, Quintile 5.	0.0310	0.0443	0.70	
Number of children 0-3 years old	0.0150**	0.0063	2.38	+++
Number of children 4-6 years old	0.0043	0.0069	0.63	+
Number of children 7-10 years old	-0.0138	0.0086	-1.61	
Number of children 11-17 years old	-0.0298***	0.0114	-2.60	-
Education, F2. (1=upper secondary school, less than 3 years)	0.1162**	0.0489	2.38	++
Education, F3. (1=post secondary school and post grad. educ.)	0.0601	0.0521	1.15	
Study allowance (1=study allowance > 5000 SEK	-0.3697***	0.0150	-24.63	---
Parental benefits (1=parental benefits > 5000 SEK	0.0002	0.0094	0.02	
Time dummy 97 (1=1997)	0.0615***	0.0034	18.05	+++
Time dummy 98 (1=1998)	0.1228***	0.0054	22.87	+++
Time dummy 99 (1=1999)	0.1953***	0.0070	28.02	+++

Table continues on the following page.

Table 6. Continues...

Sample: Male Estimation method: GMM				Sensitivity
Variables	Coefficients	Std.errors	t-value	Analysis
$\lambda_{95,96}$	-0.0034	0.0380	-0.09	
$\lambda_{96,95}$	-0.0709***	0.0212	-3.35	
$\lambda_{95,97}$	-0.0910**	0.0372	-2.45	
$\lambda_{97,95}$	-0.0411**	0.0190	-2.17	
$\lambda_{95,98}$	-0.1035***	0.0345	-3.00	
$\lambda_{98,95}$	-0.0399**	0.0176	-2.27	
$\lambda_{96,97}$	-0.0403*	0.0224	-1.80	
$\lambda_{97,96}$	0.0031	0.0182	0.17	
$\lambda_{96,98}$	-0.0302	0.0224	-1.35	
$\lambda_{98,96}$	-0.0226	0.0175	-1.29	
$\lambda_{97,98}$	-0.0086	0.0170	-0.50	
$\lambda_{98,97}$	-0.0159	0.0158	-1.01	

Wald test whether λ-terms are joint significantly different from zero;   31.8399***

Note: See *Table 1* for explanations of variables. Coefficients that are significantly different from zero at 10, 5 and 1%-level are marked with *, ** and ***. Coefficients that are positively (negatively) significantly different from zero at 10, 5 and 1%-level in the sensitivity regression are marked with + (-), ++ (--) and +++ (---). See text for explanation of the sensitivity analysis.

Note that the estimation is made in differences and the estimation of the coefficients for the variables has to rely on changes in the variables. For many individuals in the sample, education does not change at all during the period. It can also take some time to see the payoff for additional education. It is, therefore, difficult to draw conclusions from the estimated parameters for the educational variables.

6. Concluding remarks

This study empirically investigates two important effects of unemployment. Since unemployment usually means a lower disposable income, it has been suggested that a spouse could compensate for his/her partner's lower income by working more, or in some other way increasing the income. Another possible effect from unemployment is, however, an

increased risk that the couple will split up. Studies of the economic response to unemployment on the part of the spouse could be affected by a selection problem, if only couples that stay together are analyzed. Another potential problem that should be dealt with is unobserved heterogeneity. To take into account both these potential problems, this study uses a sample selection model for panel data.

The first step of the econometric model identifies characteristics that affect the probability of the couples staying together. For example, higher education and the same parents present in their own childhood seem to increase the probability that the couples will stay together. Unemployment, in particular when the income the previous year was low, seems to increase the risk for couples to split up. The purpose of modelling the first step is to construct correction terms for the earnings equation, to avoid biased estimates due to the selection problem. In this study, the first step procedure is found to be necessary for the analysis. A selection problem is indeed observed for the analyzed sample. It is also found that the variables measuring unemployment are rejected as strictly exogenous. Accordingly, a GMM-version of the sample selection model is used.

For the female sample, periods in unemployment for the spouse are seldom found to be connected to the income. However, for individuals in the fourth quintile the previous year, and who ended up in unemployment, there seems to be a compensatory behavior on the part of the female spouse. The coefficients are, in some cases, estimated with low precision, and should be interpreted with caution. The result is, however, in line with Stephens (2002) who does not find any response for low income households, while a response is found for high income households. Note, though, that this study is more detailed regarding the previous income and that no effect is found in the case of the highest income earners the previous year.

When the male sample is used, a small, but significant, effect is found for long term female unemployment when the female spouse previously was in the second quintile in the earnings distribution. When the responses are allowed to vary depending on the level of education for the male spouse, the tendency is that men with a low level of education are more likely to compensate positively. Men with a high level of education are found to have a lower income when a female spouse, who was in the first quintile the previous year, was unemployed.

The population investigated is comprised of quite young couples, and it is possible that different effects can be found for an older population. In particular, older couples tend to be more stable, and the selection problem could be different for an older population. It would, therefore, be interesting to see further studies on the subject. In the data used for the analysis, it was also not possible to identify different reasons for unemployment. It would be interesting to see if the compensatory behavior differs depending on whether the unemployment was due to a closure of a plant or due to a layoff.

The results in this study are based on a sample selection model for panel data and a sensitivity analysis. The sensitivity analysis is added as the dependent variable, in the equation that is of interest, is found to have an explanatory power in the selection equation. To my knowledge, a selection model that can handle that, does still not exists and an econometrical model of that kind need to be developed.

7. References

Askildsen, J. E., B. H. Baltagi & T. H. Holmås, (2003), Wage policy in the health care sector: a panel data analysis of nurses' labour supply, *Health Economics*, 12, 705-719.

Bakgrundsfakta till Arbetsmarknads- och utbildningsstatistiken, 2002:2, Statistics Sweden.

Becker, G. S., E. M. Landes & R. T. Michael, (1977), An Economic Analysis of Marital Instability, *Journal of Political Economy*, 85, 1141-1187.

Bingley, P. & I. Walker, (2001), Household Unemployment and the Labour Supply of Married Women, *Economica,* 68, 157-185.

Bound, J., D. A. Jaeger, & R. M. Baker, (1995), Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak, *Journal of the American Statistical Association*, Vol. 90, No. 430, 443-450.

Bowen, W. G. & T. A. Finegan, (1969), T*he Economics of Labor Force Participation*. Princeton: Princeton University Press.

Charles, K. K., & M. Stephens Jr., (2004), Job Displacement, Disability, and Divorce, *Journal of Labor Economics*, Vol. 22, issue 2, 489-522.

Dustmann, C. & M. E. Rochina-Barrachina, (2000), Selection Correction in Panel data Models: An Application to Labour Supply and Wages, *IZA Discussion Paper Series*, No. 162. Institute for the Study of Labor, Bonn.

Dynarski, M. & S. Sheffrin, (1987), Consumption and Unemployment, *Quarterly Journal of Economics,* 104, 411-428.

Greene, W. (2003), *Econometric Analysis*. 5th edition. New Jersey: Pearson Education International.

Heckman, J. (1979) Sample Selection Bias as a Specification Error, *Econometrica,* 47, 153-166.

Heckman, J. & T. MaCurdy, (1980), A Life Cycle Model of Female Labour Supply, *Review of Economic Studies,* 47, 47-74.

Kraft, K. (2001), Unemployment and the Separation of Married Couples, *Kyklos,* Vol. 54, Fasc. 1, 67-88.

Kyriazidou, E. (1997), Estimation of a panel data sample selection model, *Econometrica*, 65, 1335-1364.

Maloney, T. (1991), Unobserved Variables and the Elusive Added Worker Effect, *Economica,* 58, 173-187.

Melenberg, C. E., & A. van Soest, (2001), An Analysis of Housing Expenditure Using Semiparametric Models and Panel Data, *Journal of Econometrics*, Vol. 101, issue 1, 71-107.

Mincer, J. (1962), Labor Force Participation of Married Women, in *Aspects of Labor Economics*. Princeton: Princeton University Press.

Prieto-Rodríguez, J. & C. Rodríguez-Gutiérrez, (2003), Participation of married women in the European labor markets and the "added worker effect", *The Journal of Socio-Economics,* 32, 429-446.

Rochina-Barrachina, M. E. (1999), A New Estimator for Panel Data Sample Selection Models, *Annales d'Économie et de Statistique, 55/56, 153-181.*

Stephens, M. Jr., (2002), Worker Displacement and the Added Worker Effect, *Journal of Labor Economics*, July 2002, Vol. 20, issue 3, 504-537.

Weiss, Y. & R. J. Willis, (1997), Match quality, new information, and marital dissolution, *Journal of Labor Economics,* 15, no. 1, 293-329.