

IZA DP No. 6803

**Comparing Quasi-Experimental Designs and
Structural Models for Policy Evaluation:
The Case of a Reform of Lone Parental Welfare**

Chiara Daniela Pronzato

August 2012

Comparing Quasi-Experimental Designs and Structural Models for Policy Evaluation: The Case of a Reform of Lone Parental Welfare

Chiara Daniela Pronzato

*Universita' degli Studi di Torino, CHILD Collegio Carlo Alberto,
Statistics Norway, Dondena, ISER and IZA*

Discussion Paper No. 6803
August 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Comparing Quasi-Experimental Designs and Structural Models for Policy Evaluation: The Case of a Reform of Lone Parental Welfare*

This paper compares two different ways of doing policy evaluation: on the one hand, quasi-experimental methods (or “ex-post” evaluations) which exploit the introduction of a reform and identify its effect by comparing treated and untreated individuals; on the other hand, structural models (or “ex-ante” evaluations) which are based on economic theory and predict the effect of potential reforms by using the estimates of behavioural parameters. The comparison is carried out using an empirical case. In 1998, in Norway, a major welfare reform changed the rules of the most generous benefit for lone parents: it increased the amount of the benefit and introduced working requirements. Using a quasi-experimental evaluation approach, it is found a positive effect of the reform on lone mothers’ employment. In this paper, I estimate a static structural model of work and welfare participation decisions and compare the results using the two different approaches. Despite the differences in the assumptions I make for the two models, results are fairly comparable.

JEL Classification: I38, J22, C25

Keywords: lone mothers, in-work benefits, quasi-experimental evaluation designs, discrete choice modelling, ex-ante and ex-post evaluation

Corresponding author:

Chiara Daniela Pronzato
Collegio Carlo Alberto
Via Real Collegio 30
10024 Moncalieri (Torino)
Italy
E-mail: chiaradaniela.pronzato@unito.it

* The Norwegian Research Council has provided financial support for this project. I am grateful to Rolf Aaberge, Richard Blundell, Ugo Colombino, John Ermisch, Francesco Figari, Marco Francesconi, Magne Mogstad, Andreas Peichl, Steve Pudney, Trine Vattø, Katharina Wrohlich for their comments as well as participants to seminars at ISER, DIW Berlin, Dondena, the department of Economics of the University of Turin, the department of Economics of the Italian University of Switzerland, and participants to the XXIII ESPE conference in Seville, and the III IMA conference in Stockholm. Any error should be attributed to the author.

1 Introduction

The aim of this paper is to compare the effect of a reform estimated using a quasi-experimental evaluation design with the effect simulated with a static structural model of work and welfare participation decisions. The comparison is carried out exploiting the introduction of a welfare reform for lone parents in Norway.

Lone mothers are overrepresented among poor people in many European countries, with worrying consequences for themselves and their children. Also in Norway, which is known as a country of economic and welfare success, lone mothers were at least three times more likely to be poor than married mothers with children in the same age-range. To this aim, in 1998, a reform of lone parental welfare was undertaken. The maximum amount of the benefits was increased, working requirements were introduced, and new time limits were imposed.

Using the reform as an instrument, in a quasi-experimental setting, we can understand whether lone mothers' behaviour is influenced by public policies, without strong assumptions and referring only intuitively to the economic theory. However, we cannot distinguish the effects of the different parts of the reform, cannot understand the mechanisms, and cannot predict what kind of policy would have made lone mothers "better off". Thinking more generally, researchers can not always wait for "good" reforms to answer questions which are important from a policy point of view. And even if reforms are implemented, there are still things we would need to know in order to make our findings useful for policy makers.

This seems to be an area of research investigated only by a few papers, but necessary to give credibility to both the approaches, and to reconcile them. This is the appeal to young economists made by Keane (2006, 2010), during his keynote lecture at the Duke Conference on Structural Models in Labor, Aging and Health (2005), titled "Structural versus Atheoretic Approaches to Econometrics". He underlines the necessity of considering descriptive statistics, reduced and structural forms as well as experimental methods as complementary approaches to the study of the effects of policy changes. He encourages researchers to perform validation exercises to test the extent to which structural models

give “reasonable” predictions of the reality. The adjective “reasonable” may be still judged in a subjective way, but via multiple validation exercises consensus may be reached. Recently, as a part of the Mirrlees Review, Blundell (2012) has underlined the importance of different empirical strategies to evaluate the effects of earnings taxation on labour market decisions in order to design better tax policy reforms.

Examples are offered by Todd and Wolpin (2006), Blundell (2006), Brewer et al. (2006), Keane and Wolpin (2007) and, more recently, by Bernal and Keane (2010), Hansen and Liu (2011), Geyen et al. (2012), Thoresen and Vattø (2012). Todd and Wolpin (2006) use data from a randomized social experiment in Mexico to study and validate a dynamic behavioural model of parental decisions about fertility and child schooling. The PROGRESA is a randomized social experiment implemented by the Mexican government, in which around 500 rural villages were randomly assigned to participate or not in the program which provided payments to parents who regularly send their children to school. They estimate the behavioural model without using observations from the treated villages and predict the potential fertility and child schooling of families in untreated villages. The impact of the program predicted by the behavioural model tracks the experimental results. Keane and Wolpin (2007) adopt another approach to validate a behavioural model. They construct and estimate a dynamic structural model of female behaviour, in which work, welfare participation, marriage and fertility decisions are jointly considered. In order to validate the model, they use a “holdout sample”, a sample which differs from the sample used in the estimation and whose policy regime is well outside the support of the data. They use data from some US states to estimate the model, and from others to predict and validate the model. Bernal and Keane (2010) evaluate the effects of maternal work and childcare use on cognitive child development using a sample of single mothers in the National Longitudinal Survey of Youth. In order to take into account the selection process in work and childcare use, they develop a model of mothers’ employment and childcare decisions. To identify the model, they use exogenous variations in welfare rules and local demand conditions across States and over time. They also employ the same instrumental variables for a straight linear IV regression. The estimated effects on children’s cognitive development are very close when comparing the IV strategy and the structural model.

While the above studies construct and estimate structural models which are dynamic, there is a number of empirical works which validate static structural models with quasi-experimental evidence and which mainly look at labour market changes due to a change in the welfare. Brewer, Duncan, Shephard and Suarez (2006) estimate a static structural model of labour supply and programme participation using data from before and after the introduction of the Working Families' Tax Credit in the UK. They simulate the effect of the reform, taking into account all related changes in benefits and taxes, and compare the results with the ones obtained from other ex-ante (Blundell et al., 2000a, 2000b) and ex-post evaluations (Blundell et al., 2005; Francesconi and Van der Klaauw, 2004, 2007; Leigh, 2005; Gregg and Harkness, 2009). Blundell (2006) focuses on the effects of the Earned Income Tax Credit policies on lone mothers' working decisions, by validating a structural model of labour supply with a difference-in-difference evaluation strategy, and then finds the optimal policy, defined by a certain social welfare function. Other recent papers compare results from quasi-experimental methods and structural models exploiting the introduction of a certain reform: Geyen, Hann, and Wrohlich (2012) estimate the introduction of a parental leave reform in Germany by comparing working behaviour of mothers of children born just before or after the reform and compare results with the ones obtained by a structural model of return to work; Thoresen and Vattø (2012) evaluates a tax reform by comparing a before-after change in labour supply with the effect predicted by a structural model of labour supply; Hansen and Liu (2011) compares the effect of an increase of the generosity of welfare benefits for young people in Quebec predicted by their structural model of labour supply and welfare participation with the one estimated with a discontinuity regression model by Le Mieux and Milligan (2008).

In this paper, I compare the effect of the 1998 Norwegian welfare reform on lone mothers' working decisions adopting two different empirical strategies but using the same data and the same outcome variable for the two analyses. First, I estimate the effect of the reform by employing a quasi-experimental evaluation design. Second, I estimate a static structural model of work and welfare participation decisions from which I simulate the effect of the reform. The structural model, together with the econometric specification, is explained in Section 2. The data and the procedure to derive the budget sets are presented in Section 3.

After estimating the structural parameters (Section 4), the reform is simulated, and its predictions compared with the estimated effects obtained with the quasi-experimental method: despite the model making strong assumptions, the structural results are close to the quasi-experimental ones (Section 5). Sensitivity analyses (Section 6) and conclusions (Section 7) follow.

2 A Model of Work and Welfare Participation Decisions ¹

The Norwegian register data, I use for the estimation of the model, provide accurate information on incomes and demographic characteristics but not hours of work. The model outlined below takes this feature into account, allowing time of work to be measured with an error.

A lone mother, labelled n , is assumed to maximize a utility function

$$U_n(x, t, w) \tag{1}$$

under the budget constraint

$$x_n = T(l_n, t_n, y_n) \tag{2}$$

where

- x_n is the net household income,
- l_n is the gross monthly labour income of the lone mother in a full time job,
- t_n is the number of equivalent full time months of work in one calendar year,
- w_n is a welfare participation indicator,
- y_n is exogenous household gross income,

¹ This paragraph follows Train's book on Discrete Choice Methods and Simulations, chapter 2 (2003). Other papers used as references to write the model are Mc Fadden (1974), Moffitt (1983), MaCurdy et al. (1990), Ilmakunnas and Pudney (1990), Van Soest (1995), Hoynes (1996), Aaberge et al. (1999), Creedy and Kalb (2005), Creedy et al. (2006), Keane (2011), Bargain et al. (2012).

- $T(.)$ is the tax-benefit function which transforms gross income into net income.

The lone mother faces a set of J discrete alternatives, defined by the combination of earnings and welfare participation decisions. She knows how much utility she would get from each alternative j and chooses the alternative which provides the largest one. We can decompose the utility function into two parts: the deterministic (or systematic) part and the stochastic part

$$U_{nj} = V_{nj} + \varepsilon_{nj} \quad \forall j \in J \quad (3),$$

where V_{nj} captures the portion of utility which derives from observable characteristics, while ε_{nj} the portion from unobservable ones. The deterministic part of the utility V_{nj} may be seen as a function which relates the observable characteristics to the lone mothers' utility

$$V_{nj} = V(z_{nj}, s_n) \quad \forall j \in J \quad (4),$$

where z_{nj} are the observed attributes of the alternatives as faced by the lone mother, and s_n the observed socio-demographic characteristics of the lone mother. I specify the deterministic part of utility to be linear in parameters with a constant

$$V_{nj} = q'_{nj} \theta + k_j \quad \forall j \in J \quad (5)$$

where q_{nj} is a vector of variables that relate to alternative j as faced by the lone mother n , θ are the coefficients of these variables, and k_j is a constant that is specific to alternative j . The constant k_j captures the average effect on utility of all factors not included in the model. The vector z_{nj} includes the net income available to the lone mother at alternative j and its square, the time of work required by alternative j and its square, a welfare participation indicator, and their interactions. The socio-demographic variables s_n cannot enter the model directly, since they do not vary across alternatives. They are interacted with net income, time of work and the welfare indicator to allow utility from income and disutility from time

of work and welfare participation to be different for women with different levels of education, age, nationality, numbers and ages of children:

$$V_{nj} = \beta_1 x_{nj} + \beta_2 x_{nj}^2 + \beta_3 t_{nj} + \beta_4 t_{nj}^2 + \beta_5 w_{nj} + \beta_6 x_{nj} t_{nj} + \beta_7 t_{nj} w_{nj} + \beta_8 x_{nj} w_{nj} + k_j + (x_{nj} s_n)' \delta + (t_{nj} s_n)' \gamma + (w_{nj} s_n)' \lambda \quad \forall j \in J \quad (6)$$

where x_{nj} is her net household income, t_{nj} her time of work, w_{nj} her welfare participation indicator in each alternative j , and s_n are her demographic characteristics.

Time of work t_n is not observed in the register data. I derive the expected time of work \bar{t}_n , expressed in equivalent full time months of work in a year, as the ratio between each woman's annual earnings in the register data ($l_n t_n$) and the predicted monthly earnings from survey data (\bar{l}_n) in a full time job of a woman with same human capital characteristics:

$$\bar{t}_n = \frac{l_n t_n}{\bar{l}_n} \quad (7).$$

The relationship between true time of work t_n and expected time of work \bar{t}_n is given by

$$t_n = \frac{\bar{l}_n}{l_n} \bar{t}_n = \alpha \bar{t}_n \quad (8),$$

where α is negatively correlated with the unobservable characteristics which make a woman earn more. If a woman earns more than what, on average, a woman with the same observable characteristics does, it means that she needs to work less time than what I predict as expected time of work. α connects true and expected time of work and is assumed to be normally distributed. Therefore (6) becomes

$$\begin{aligned}
V_{nj} = & \beta_1 x_{nj} + \beta_2 x_{nj}^2 + \tilde{\beta}_3 \bar{t}_{nj} + \tilde{\beta}_4 \bar{t}_{nj}^2 + \beta_5 w_{nj} + \tilde{\beta}_6 x_{nj} \bar{t}_{nj} + \tilde{\beta}_7 \bar{t}_{nj} w_{nj} + \beta_8 x_{nj} w_{nj} + \\
& + k_j + (x_{nj} s_n)' \delta + (\bar{t}_{nj} s_n)' \tilde{\gamma} + (w_{nj} s_n)' \lambda \quad \forall j \in J
\end{aligned} \tag{9}$$

where, for example,

$$\tilde{\beta}_6 = \beta_6 \alpha \tag{10}.$$

The model I estimate allows disutility from time $\tilde{\beta}_3$ to be different for women with different unobservable characteristics:

$$\begin{aligned}
V_{nj} = & \beta_1 x_{nj} + \beta_2 x_{nj}^2 + (\beta_3 \nu) \bar{t}_{nj} + \tilde{\beta}_4 \bar{t}_{nj}^2 + \beta_5 w_{nj} + \tilde{\beta}_6 x_{nj} \bar{t}_{nj} + \tilde{\beta}_7 \bar{t}_{nj} w_{nj} \\
& + \beta_8 x_{nj} w_{nj} + k_j + (x_{nj} s_n)' \delta + (\bar{t}_{nj} s_n)' \tilde{\gamma} + (w_{nj} s_n)' \lambda \quad \forall j \in J
\end{aligned} \tag{11}.$$

Some points need clarification. First, the method of measuring expected time of work, by comparing observed earnings and predicted earnings, holds if the wage rate is constant over time of work. Second, ν coincides with α only if there is no difference in tastes due to unobservables among women. However, I do not need to identify α because the main aim is to take into account that time of work is measured with an error. Third, $\tilde{\beta}_4, \tilde{\beta}_6, \tilde{\beta}_7, \tilde{\gamma}$ should be also allowed to vary among women but, in practice, the model does not converge when allowing unobservable heterogeneity in many parameters. Fourth, the estimated coefficients $\tilde{\gamma}$ related to the interactions between time of work and variables also used in the earnings equation (age and level of education) capture different disutility from time of work but can also capture misspecification of the earnings equation, so that their interpretation is not possible.

The probability that a lone mother chooses the alternative i is

$$P_{ni} = P(U_{ni} > U_{nj}) \quad \forall j \neq i$$

$$\begin{aligned}
&= P(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}) \quad \forall j \neq i \\
&= P(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}) \quad \forall j \neq i
\end{aligned} \tag{12}.$$

The stochastic component ε_{nj} is assumed to be iid extreme value for all j . The probability that the lone mother n chooses i is then given by

$$P_{ni} = \frac{e^{\beta'z_{ni}}}{\sum_j e^{\beta'z_{nj}}} \tag{13},$$

which is the expression of the conditional logit regression.

3 The Data

The data used for the empirical analysis are from the register data of the Norwegian population in the period 1993-2001.² The sample for the quasi-experimental evaluation design is composed of 1,121,898 women with children aged between 4 and 9 years old³: lone mothers (treatment group) and mothers in a couple (control group), before and after the reform.⁴ The sample for the structural model is composed of the sub-sample of lone mothers observed before the reform (7,921), representing what one would use in a typical ex-ante evaluation.

² The register panel data set with household and demographic information are merged with detailed income data from the Tax Assessment Files through unique individual identifiers. The income data are collected from tax records and other administrative registers rather than interviews and self-assessment methods. The coverage and reliability of Norwegian register data are considered to be exceptional, as is documented by the fact that the quality of such national data sets received the highest rating in a data quality survey in the Luxembourg Income Study database (Atkinson et al., 1995).

³ The reform of the transitional benefit involved all lone mothers with children younger than 10 years old but working requirements were only introduced for lone mothers with the youngest child older than 3 (see Mogstad and Pronzato 2008, 2012).

⁴ Students, self-employed, as well as individuals receiving permanent disability benefits are dropped from the analysis (see Mogstad and Pronzato 2008, 2012).

I assume lone mothers face at most 8 alternative choices, given by the joint decision of how much to work (4 alternatives) and whether or not participating in the welfare (2 alternatives).

As explained in Section 2, expected time of work is obtained comparing annual earnings observed in the register data with potential monthly full time earnings from survey data. In order to construct potential earnings, I use the Norwegian part of the European Union Survey of Income and Living Conditions for the year 2004, I select women in the same age-range (18-55), and I estimate a Heckman regression. The dependent variable is hourly gross labour income. In the outcome equation I include two dummy variables for education (secondary and tertiary education), a variable for potential working experience (age - years of schooling - 7), its square, and a part time dummy.⁵ In the selection equation, I also consider the presence of dependent children, other household income, whether being in a couple, and living in a city. Results are reported in Table A1. In order to make survey earnings comparable to earnings in the register data, predicted hourly earnings are multiplied by typical hours of work in a full time job (38) and number of weeks in a month, and adjusted in order to take into account nominal and real growth.⁶

The 4 work alternatives are defined in the following way:

- 1) First work alternative (which will be called “no work”): ratio between annual observed earnings and expected monthly earnings in a full time job smaller than 3, which corresponds to less than 9.5 hours a week (on average, in the data, 1 hour and half per week).
- 2) Second work alternative (which will be called “short part time”): ratio between observed annual earnings and expected monthly earnings in a full time job larger or equal to 3 and smaller than 6, which corresponds to 9.5-19 hours a week (on average, in the data, 13 hours per week).

⁵ I include a part time dummy to test whether the wage rate can be considered constant over time of work. Part time wage rate is not significantly different from full time wage rate, as shown in Table A1.

⁶ Prices are deflated to €1998. Real growth is taken into account looking at the variation, year by year, of the basic amount (*grunnbeløp*), which is the official reference amount used for the up-rating of benefits and pensions.

- 3) Third work alternative (which will be called “part time”): ratio between observed annual earnings and expected monthly earnings in a full time job larger or equal to 6 and smaller than 9, which corresponds to 19-28.5 hours a week (on average, in the data, 22 hours per week).
- 4) Fourth work alternative (which will be called “full time”): ratio between observed annual earnings and expected monthly earnings in a full time job larger or equal to 9, which corresponds to more than 28.5 hours a week (on average, in the data, 33 hours per week).

In the observed choice, the three objects of the utility function are defined as follows: (i) the observed welfare participation decision, (ii) the net income which derives from observed earnings through the tax-benefit function (2) and (iii) the expected number of months of work, obtained dividing the observed annual earning by potential monthly earnings in a full time job. For the other 7 alternatives I construct counterfactuals.

Suppose her observed earnings are €7,500 and she participates in the welfare (see example in Table 1). Given her human capital characteristics, she is supposed for example to earn €2,500 per month in a full time job. I classify her as working “part time” ($17,500 / 2,500 = 7$ equivalent full time months; 22 hours per week). I construct three other earning alternatives: “no work”, working “short part time”, working “full time” (see Table 1, first five columns). The number of months chosen for each untaken work alternative (no work / short part time / full time) is drawn from the distribution of months of people choosing that alternative (no work / short part time / full time) (Aaberge et al., 2009). Predicted earnings are then imputed. In the example, Table 1, the drawn numbers of months are 0, 4 and 12, and earnings are, respectively, €0, €10,000, and €30,000.

For each earning alternative, she can decide whether to participate in the welfare. The transitional benefit is calculated as follows. The maximum annual amount of the benefit is around €8,000 per year. From this maximum amount, 40% of earnings exceeding €2,500 are subtracted. In Table 1, 6th column, we can see the corresponding amounts. For this woman, the 7th alternative is dropped, since the related full time earnings are too large to be still eligible for the benefit.

I then simulate the childcare benefit, another benefit which depends on labour supply, given as a reimbursement for extra-costs for childcare, occurred when the mother works.

All other remaining benefits are only available in the data as a total amount. However, none of them should formally depend on her working decisions. In order to verify this, I regress the total amount of the remaining benefits on the number and age of children, being an immigrant, time of work, and earnings. From Table A2, we observe that part of these benefits seems indeed to be related to her working decisions and her earnings: not working (rather than working full time) increases the average amount of benefits of 1,235€ per year; an increase of 1,000€ in her earnings decreases the average amount of 44€. Nevertheless, it represents a small part of the benefits: at most – for a non-working woman – it is 14% of the total amount of these benefits. I am probably capturing programs like the social assistance, whose rules do not take into account directly her working decisions, but which are related to situations of poverty and exclusion. I use the estimated coefficients of this regression to simulate this part of welfare.

Finally, I simulate taxes, and obtain the total net income she can have in different work/welfare alternatives (8th column, Table 1). The variable, in the 9th column (Table 1), indicates the decision taken and observed in the data.

Descriptive statistics of the sample are shown in Table 2.

4 Model Estimates

I estimate the effects of income, time of work, welfare participation and their interactions with other socio-demographic variables, on the probability of choosing one of the alternatives, using a mixed logit specification with the coefficient of time of work treated as random coefficient, assumed to be normally distributed, as outlined in Section 2. Results are reported in Table 3.

The model is estimated without any restriction imposed on the utility function. To check that the utility function respects the concavity and monotonicity properties, I check the derivatives with respect to the utility arguments. The first derivative with respect to income is positive for the whole sample as well as the first derivative with respect to time of work

is negative for the whole sample. Second derivatives are in the expected direction, as shown in Table 3. Utility is decreasing in welfare participation for 96% of the sample.

The standard deviation of the random coefficient is significantly different from zero, revealing an important role of unobserved heterogeneity and/or measurement error.

The interaction between income and time of work is positive, which may be explained by the presence of better positions in the labour market that, even if imply longer hours of work, increase woman's utility. The interaction between welfare participation and income is positive: since a large part of lone mothers' income comes from other benefits in Norway, the positive interaction could reveal that the cost of participating is lower for women who also participate in other welfare programs. The interaction between welfare and time of work is also positive: women who work more may be more informed because they are more likely to talk with other people at the place of work or they may suffer less from welfare stigma because they feel they do not completely depend on welfare. In the Norwegian context, where applications can be done on line and transfers can be received in the bank account without friends and family necessarily knowing, we could expect the role of the welfare stigma to be relatively less important.

Results concerning number and age of children are in the expected direction: on the one hand, having more and younger children increases the cost of working; on the other hand, it increases utility from income.

Immigrant women have more disutility from time of work. This finding could result from the fact that, given their level of education, they are in poorly paid jobs.

Younger women have less disutility from participating in the welfare while the cost of the welfare is not linear by years of education. Women with secondary schooling have less disutility from participating in the welfare than higher and lower educated women. This may capture different aspects of welfare participation: on the one hand, if information is needed then better educated women may be more prompt to apply for the benefit; on the other hand, better educated women may suffer more to be dependent from welfare.

5 Comparing the Estimated Effects of the Reform

In this section, I compare the estimated effects of the reform on lone mothers' earnings from the quasi-experimental evaluation design with the simulated effects from the structural model. In the quasi-experimental setting, the effect of the reform is estimated using a sample of lone mothers and mothers in a couple, before and after the reform, and employing a triple-difference estimator (Mogstad and Pronzato, 2008 and 2012).⁷ Results are shown in the top part of Table 4 for all women, and by level of education, together with the 95% confidence intervals: we observe that the reform increases significantly earnings of low and medium educated lone mothers.

In order to simulate the reform with the structural model, I need to parameterize the transitional benefit according to the new rules. There are three important changes. First, working requirements are imposed. Second, the age limit for eligibility on the youngest child is lowered, and time limits on welfare participation are introduced. And third, in-work benefit levels are raised.

The increase of the maximum amount (from around 8,300€ to 9,500€ per year) not only makes the transitional benefit more generous but also makes women more likely to be eligible: before the reform, only women earning less than €1,900 per month can receive the benefit while, after the reform, women earning until €2,200 per month are also eligible. This results in a larger number of alternatives in the choice set for those women now eligible to receive the benefit. According to the change of the age limit, women with the youngest child aged 9 years old (depending on the months of birth) are not allowed to receive the transitional benefit anymore. The reform requires lone mothers to be in training, to work at least part time, or to seek work. The law does not give details about the

⁷ The effect of the reform is not estimated using an evaluation design which exploits observations for the same women before and after the reform, as the traditional difference-in-difference estimator, since the reform is characterized by a long phase-in period: all lone mothers who were already in welfare at the time of the reform were allowed to receive the benefit according to the pre-reform rules for another 3 years. Therefore, we evaluate the effect of the reform by comparing the effect of becoming a lone mother on earnings in the pre-reform and post-reform period. We assume that, in absence of the reform, mothers in a couple who become lone mothers after the reform would behave the same as mothers in a couple who became lone mothers before the reform. Considering only the flow of new lone mothers, we overcome the phase-in problem (see Mogstad and Pronzato 2008, 2012).

“training” and “seeking work” activities. I do not have any information in what these activities consist of, whether it was difficult to have a training period, what women were asked in case they were seeking work. And I do not know whether these activities were easily approved by the public government, and for how long they were compatible with being eligible for the benefit. Moreover, register data do not have information on training and periods of seeking work, so that I cannot know who was taking this decisions, even after the reform. What I can do is to assume that lone mothers receiving the maximum amount of the transitional benefit after the reform are engaged in one of the two activities. In fact, the maximum amount is given only to women earning less 200€per month and it is reasonable to assume there is no part job in Norway paid less than 200€ a month. The percentage of non-working women on welfare after the reform (supposed to be in training or seeking-work activities) is 7.0% while the percentage of non-working women on welfare is 18.5% before the reform. While before the reform, the possibility of receiving the transitional benefit and not working was a woman’s decision, after the reform it is the result of the woman’s decision and the new constraints imposed by the law. We may expect women with lower level of education to be more likely to be observed in training or seeking work activities after the reform, but this does not seem to be the case when looking at post-reform data. The percentage of non-working women receiving the benefit before the reform was 26.4% among low educated women, 14.0% among medium educated women, 11.0% among highly educated women. The percentage of non-working women receiving the benefit after the reform (assumed to be in training or seeking work activities) is 8.7% among low educated women, 6.3% among medium educated women, 7.6% among low educated women. In order to reproduce what I observe in the data, I drop randomly the alternatives of non-working and participating in the welfare for a number of women so that the percentage of women taking this decision - with the new rules - is 7.0%⁸. The simulated

⁸ Another possibility would be to compare women participating in the welfare, before and after the reform, by time of work. We would observe – after the reform - a significant reduction of non-working women receiving the benefit, and a very small reduction of women working short-time receiving the benefit. However, these calculations would be based on a measure of time of work which is affected by measurement error and which depends strongly on the estimated earnings equation. Nevertheless, results would be very close to the experimental ones (*Comparing Quasi-Experimental Designs and Structural Models for Policy Evaluation*:

effects of the reform from the structural model are calculated as weighted earnings, given by the sum of earnings in each alternative times the probability of choosing that alternative. Results are shown in the bottom part of Table 4.

Table 4 shows that the results given by the “structural” simulated effects and the “experimental” estimated effects are very close: the quasi-experimental design suggests an increase in earnings of €384 while the structural predictions indicate an increase of 450€⁹. The structural model predicts well also by level of education: the positive effect is larger for low-medium educated women, while slightly negative for highly educated women. Imposing the proportion of women in training or seeking work activities to 7.0% implies, according to the structural estimates, to grant the possibility of receiving the benefit without working to 40 women out of 100.

At this point, we can separate the effects of the introduction of the working requirements and new age limits from the effect of more generous benefits. The bottom part of Table 4 summarizes the results. The introduction of working requirements has increased women’s earnings, as expected. The effect is larger for low and medium educated women than for highly educated women. Also the new age limit has a positive but small effect on work decisions. Making the benefit more generous has the expected negative effect on annual earnings. The effect is relatively large also for highly educated women. In fact, for highly educated women, the increase in the maximum amount has made them eligible in more work alternatives.

6 Sensitivity Analyses

In this section, I repeat the analyses, using two alternative specifications of the econometric model which are commonly used in this strand of the literature. I then compare the

The Case of a Reform of Lone Parental Welfare in When Women Work, PhD Thesis, ISER University of Essex 2009).

⁹ In order to calculate the confidence intervals around the predictions from the structural model, I employ the bootstrap method: I draw 100 new samples from the original one, each of them containing the original number of observations ($N = 7,921$), where each observation may be repeated more than once (with replacement); I re-estimate the model using each of the 100 new samples; I parameterize the reform and get 100 sets of predictions. From these predictions I calculate means, standard errors, and confidence intervals.

predictions of each model with the ones provided by the main specification outlined in Section 2 and 3. Table 5 summarizes the results. Each panel reports the results of a different model: the predicted effects of the increase of the benefit, of the new age limit, of the working requirements and of the whole effect of the reform. The top panel reports results when using the main model (effect of time of work as random coefficient).

I first estimate the model without unobserved heterogeneity (2nd panel, Table 5). The effect of the reform would be larger, driven by a larger effect of the introduction of the working requirements on earnings. Controlling for unobserved heterogeneity seems to be important not to overestimate the impact of the working requirements.

I then estimate the model allowing unobserved heterogeneity in time of work, income and welfare participation (3rd panel, Table 5). Results are almost the same. In fact, the standard deviation of the random coefficients related to income and welfare are not significantly different from zero.

7 Conclusions

In this paper, I compare the effect of the 1998 Norwegian welfare reform on lone mothers' earnings estimated using a quasi-experimental evaluation design and a static structural model of work and welfare participation decisions. The reform increases the maximum amount of the transitional benefit, introduces new working requirements and changes time limits in order to be eligible for it. From both the evaluation methods, we observe a positive effect on lone mothers' earnings, driven by behavioural responses of lower and medium educated women.

The paper contributes to the policy evaluation literature by offering an example of evaluation of the same reform through different empirical strategies. It can contribute to the policy debate by supplying a validated structural model to design policies which aim at increase labour market participation or reduce poverty, providing also an estimate of the government expenditure.

The two strategies help the understanding of the policy impact in a complementary way: while the focus of the quasi-experimental evaluation design is to measure what really

happened, the challenge of the structural model is to predict what potentially can happen. Both aspects are important from a policy point of view. The fact that predictions provided by the structural model track the results of the quasi-experimental evaluation gives credibility to both the approaches. From the researcher's point of view, working with the two evaluations methods helps one not to forget part of the story: while the attention when working with the structural model is to understand the mechanisms and to carefully reproduce what opportunities individuals face, all economic predictions need to be compared with a measure of what happened. For example, if only working with the quasi-experimental method, I would have paid less attention to the take-up decision: despite the generosity of the welfare, only 70% of the women take up the transitional benefit at the time of the reform. This may be important when judging the strength of the effects of the reform or, for example, when designing a policy which aims at minimizing poverty. On the other hand, if working only with the structural model, I would have not carefully analyzed how new working requirements have been implemented at the time of the reform.

This study obviously carries with it a number of limitations mainly due to the lack of information. From a substantive point of view, I would have liked to have more information on what opportunities women faced at the time of the reform. How easy/difficult/demanding it was for them to be "seeking work", to find any "training" to attend, and to have these activities approved? For how long these activities were compatible with being eligible? Another further improvement could have derived from using a validated micro-simulation model of taxes and benefits for Norway.

References

- Aaberge R., Colombino U., and Strom S. (1999), *Labour Supply in Italy: an Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints*, Journal of Applied Econometrics, 14: 403-422.
- Aaberge R., Colombino U., and Wennemo T. (2009), *Evaluating Alternative Representations of the Choice Sets in Models of Labour Supply*, Journal of Economic Surveys, 23(3): 586-612.
- Atkinson A., Rainwater L., and Smeeding T. (1995): *Income Distribution in OECD Countries*, Paris, OECD.
- Bargain O., Orsini K., and Peichl A. (2012), *Comparing Labor Supply Elasticities in Europe and the US: New Results*, IZA Discussion Paper 6735.
- Bernal R. and Keane M. P. (2010), *Quasi-structural estimation of a model of childcare choices and child cognitive ability production*, Journal of Econometrics, 156: 164-189.
- Blundell R. (2006), *Earned Income Tax Credit Policies: Impact and Optimality. The Adam Smith Lecture 2005*, Labour Economics, 13: 423-443
- Blundell R. (2012), *Tax Policy Reform: The Role of Empirical Evidence*, Journal of the European Economic Association, 10(1): 43-77.
- Blundell R., Brewer M., and Shephard A. (2005), *Evaluating the Labour Market Impact of Working Families Tax Credit using Difference-in-differences*, Open Access from University College London, 18451.
- Blundell R., Duncan A., McCrae J., and Meghir C. (2000a), *Evaluating In-Work Benefit Reform: the Working Families Tax Credit in the UK*, JCPR Working Paper no 160.
- Blundell R., Duncan A., McCrae J., and Meghir C. (2000b), *The Labour Market Impact of the Working Families' Tax Credit*, Fiscal Studies, 21: 75-104.
- Brewer M., Duncan A., Shephard A., and Suarez M. J. (2006), *Did Working Families' Tax Credit Work? The Impact of In-work Support on Labour Supply in Great Britain*, Labour Economics, 2006: 699-720.
- Creedy J. and Kalb G., (2005), *Discrete Hours Labour Supply Modelling: Specification, Estimation and Simulation*, Journal of Economic Surveys, 19: 697-738.
- Creedy J., Kalb G., and Scutella R. (2006), *Income Distribution in Discrete Hours Behavioural Microsimulation Models: An Illustration*, Journal of Economic Inequality, 4: 57-76.
- Francesconi M. and Van der Klaauw W. (2004), *The Consequences of 'In-work' Benefit Reform in Britain: New Evidence from Panel Data*, ISER Working Paper no 13.
- Francesconi M. and Van der Klaauw W. (2007), *The Socioeconomic Consequences of 'In-work' Benefit Reform for British Lone Mothers*, Journal of Human Resources, 42(1): 1-31.
- Geyer J., Hann P. and Wrohlich K. (2012), *Labor supply of mothers with young children: Validating a structural model using a natural experimental*, mimeo.
- Gregg P. and Harkness S. (2009), *Welfare Reform and Lone Parents in the UK*, Economic Journal, 119(535): F38-F65.

- Hansen J. and Liu X. (2011), *Estimating Labor Supply Responses and Welfare Participation: Using a Natural Experiment to Validate a Structural Labor Supply Model*, IZA Discussion Paper 5718.
- Hoynes H. V. (1996), *Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP*, *Econometrica*, 64: 295-332.
- Ilmakunnas S. and Pudney S. (1990), *A Model of Female Labour Supply in the Presence of Hours Restrictions*, *Journal of Public Economics*, 41: 183-210.
- Keane M. P. (2006), *Structural vs. Atheoretic Approaches to Econometrics*, Keynote Address at the Duke Conference on Structural Models in Labor, Aging and Health 2005, mimeo.
- Keane M. P. (2010), *Structural vs. Atheoretic Approaches to Econometrics*, *Journal of Econometrics*, 156(1): 3-20.
- Keane M. P. (2011), *Labour Supply and Taxes: A Survey*, *Journal of Economic Literature*, 49(4): 961-1075.
- Keane M. P. and Wolpin K. I. (2007), *Exploring the Usefulness of a Nonrandom Holdout Sample for Model Validation: Welfare Effects on Female Behaviour*, *International Economic Review*, 48: 1351-1378.
- Leigh A. (2005), *Optimal Design of Earned Income Tax Credits: Evidence from a British Natural Experiment*, CEPR Discussion Paper no 488.
- Lemieux T. and Milligan K. (2008), *Incentive Effects of Social Assistance: a Regression Discontinuity Approach*, *Journal of Econometrics*, 142: 807-828.
- MaCurdy T., Green D., and Paarsch H. (1990), *Assessing Empirical Approaches for Analyzing Taxes and Labor Supply*, *The Journal of Human Resources*, 25: 415-490.
- McFadden, D. (1974), *Conditional Logit Analysis of Qualitative Choice Behavior*, in *Frontiers in Econometrics*, ed. P. Zerembka, 105-142. New York: Academic Press.
- Moffitt R. (1983), *An Economic Model of Welfare Stigma*, *The American Economic Review*, 73: 1023-1035.
- Mogstad, M. and C. Pronzato (2008), *Are Lone Mothers Responsive to Policy Changes? The Effects of a Norwegian Workfare Reform on Earnings, Education, and Poverty*, *Statistics Norway Discussion Paper*, no 533.
- Mogstad, M. and C. Pronzato (2012), *Are Lone Mothers Responsive to Policy Changes? Evidence from a Workfare Reform in a Generous Welfare State*, *Scandinavian Journal of Economics*, forthcoming.
- Pronzato C. (2009), *Comparing Quasi-Experimental Designs and Structural Models for Policy Evaluation: The Case of a Reform of Lone Parental Welfare in When Women Work*, PhD Thesis, ISER University of Essex.
- Thoresen T. O. and Vattø T. E. (2012), *Income Responses to Tax Changes. Reconciling Results of Quasi-Experimental Evaluation and Structural Labor Supply Model Simulations*, mimeo.
- Todd P. E. and Wolpin K. (2006), *Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioural Model of Child Schooling and Fertility*, *The American Economic Review*, 96: 1384-1417.
- Train K. E. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press.
- Van Soest A. (1995), *Structural Models of Family Labor Supply. A Discrete Choice Approach*, *The Journal of Human Resources*, 30: 63-88.

Tables

Table 1: An example of Choice set

| Work | Take up | Alternative | Time of work | Labour income | Transitional benefit | ... | Total net income | Decision |
|------------|---------|-------------|--------------|---------------|----------------------|-----|------------------|----------|
| No Work | Yes | 1 | 0 | 0 | 8,000 | | 17,000 | 0 |
| No Work | No | 2 | 0 | 0 | 0 | | 9,000 | 0 |
| Short Part | Yes | 3 | 4 | 10,000 | 5,000 | | 23,000 | 0 |
| Short Part | No | 4 | 4 | 10,000 | 0 | | 18,000 | 0 |
| Part Time | Yes | 5 | 7 | 17,500 | 2,000 | | 26,000 | 1 |
| Part Time | No | 6 | 7 | 17,500 | 0 | | 24,000 | 0 |
| Full Time | Yes | 7 | 12 | 30,000 | 0 | | - | - |
| Full Time | No | 8 | 12 | 30,000 | 0 | | 33,000 | 0 |

Notes: The choice set of a woman who takes-up the transitional benefit with observed earnings equal to €17,500 and potential monthly earnings equal to €2,500.

Table 2: Descriptive statistics

| Variable | Mean |
|--------------------------------|-------------------|
| Income | 24,855 (6,550) |
| Time of work | 8.38 (5.14) |
| Welfare participation | 0.489 |
| Age: younger than 32 years old | 0.362 |
| Age: 32-36 years old | 0.341 |
| Age: older than 36 years old | 0.297 |
| < 11 years of schooling | 0.253 |
| 11-12 years of schooling | 0.553 |
| > 12 years of schooling | 0.194 |
| Youngest child 4-5 years old | 0.499 |
| Youngest child 6-7 years old | 0.310 |
| Youngest child 8-9 years old | 0.191 |
| One child | 0.320 |
| Two children | 0.455 |
| More than two children | 0.225 |
| Immigrant | 0.010 |
| Observations | 7,921 |

Notes: average values, standard deviations in brackets for continuous variables. Time of work expressed in equivalent full time months.

Table 3: Model Estimates

| | Beta | St err | | Beta | St err | | Beta | St err |
|-------------------------------|-----------|--------|------------------------------|-----------|--------|--------------------------------|-----------|--------|
| Income | 1.090*** | 0.074 | Time | -0.793*** | 0.116 | Welfare | -3.647*** | 0.550 |
| | | | St dev (time) | 0.254*** | 0.036 | | | |
| Income sq. | -0.012*** | 0.001 | Time sq. | 0.001 | 0.003 | | | |
| Income*time | 0.009*** | 0.003 | Time*welfare | 0.169*** | 0.046 | Income*welfare | 0.031*** | 0.011 |
| Income interacted with | | | Time interacted with | | | Welfare interacted with | | |
| Mother's age (< 32) | 0.007 | 0.030 | Mother's age (< 32) | -0.003 | 0.038 | Mother's age (< 32) | 0.393** | 0.156 |
| Mother's age (32-36) | -0.001 | 0.025 | Mother's age (32-36) | 0.053 | 0.034 | Mother's age (32-36) | 0.342*** | 0.124 |
| Mother's age (> 36) | | | Mother's age (> 36) | | | Mother's age (> 36) | | |
| Schooling (< 11) | 0.077* | 0.042 | Schooling (< 11) | -0.250*** | 0.060 | Schooling (< 11) | -0.213 | 0.224 |
| Schooling (11-12) | -0.036 | 0.035 | Schooling (11-12) | -0.128** | 0.054 | Schooling (11-12) | 0.547*** | 0.158 |
| Schooling (> 12) | | | Schooling (> 12) | | | Schooling (> 12) | | |
| One child | -0.154*** | 0.031 | One child | 0.389*** | 0.044 | One child | 0.044 | 0.150 |
| Two children | -0.045* | 0.024 | Two children | 0.157*** | 0.033 | Two children | -0.049 | 0.123 |
| More than two | | | More than two | | | More than two | | |
| Youngest child 4-5 | 0.068** | 0.027 | Youngest child 4-5 | -0.133*** | 0.036 | Youngest child 4-5 | -0.208* | 0.126 |
| Youngest child 6-7 | 0.046* | 0.027 | Youngest child 6-7 | -0.081** | 0.036 | Youngest child 6-7 | -0.003 | 0.12 |
| Youngest child 8-9 | | | Youngest child 8-9 | | | Youngest child 8-9 | | |
| Immigrant | 0.038 | 0.082 | Immigrant | -0.280** | 0.113 | Immigrant | -0.261 | 0.496 |
| No-work intercepts | | | Short time intercepts | | | Part time intercepts | | |
| Welfare | -1.064*** | 0.242 | Welfare | -1.421*** | 0.159 | Welfare | -0.606*** | 0.106 |
| No welfare | 0.654** | 0.323 | No welfare | -1.059*** | 0.207 | No welfare | -0.770*** | 0.120 |
| Observations | 59,293 | | | | | | | |

Notes: Mixed logit regression with time treated as random coefficient; *** significant at 1% level, ** significant at 5% level, *significant at 10% level. Income variables divided by 1,000. Time of work expressed in equivalent full time months.

Table 4: Comparison of the Effects of the Reform on Earnings, obtained with the Quasi-Experimental Design and the Structural Model

| | All women | Low educated (< 11 years of schooling) | Medium educated (11-12 years of schooling) | High educated (>12 years of schooling) |
|---------------------------------------|-----------|--|--|--|
| Quasi – experimental design | | | | |
| Estimated effect | 384 | 658 | 476 | -79 |
| 95% confidence interval | (195,572) | (276,1040) | (238,714) | (-550,392) |
| Observations | 1,121,898 | 207,808 | 633,870 | 358,294 |
| Average earnings before the reform | 16,748 | 11,692 | 15,982 | 25,508 |
| Structural model | | | | |
| Reform | 450 | 788 | 516 | -176 |
| 95% confidence interval | (399,501) | (728,848) | (462,569) | (-215,-136) |
| Increased generosity | -835 | -907 | -861 | -665 |
| Age limit | 115 | 15 | 156 | 130 |
| Activity requirements | 916 | 1,071 | 1,030 | 390 |
| Observations | 7,921 | 2,002 | 4,380 | 1,539 |

Notes: the top panel reports the estimated effects of the reform when using the quasi-experimental design, together with the 95% confidence intervals, the number of observations, and the average earnings before the reform; the bottom panel reports the simulated effects of the reform when using the structural model, together with the 95% confidence intervals, the simulated effects of each policy parameter changed by the reform, and the number of observations.

Table 5: Sensitivity Analyses

| | All women | Low educated | Medium educated | High educated |
|---|------------------|---------------------|------------------------|----------------------|
| Random coefficients: | | | | |
| Time (Table 3) | | | | |
| Reform | 450 | 788 | 516 | -176 |
| Increased generosity | -835 | -907 | -861 | -665 |
| Age limit | 115 | 15 | 156 | 130 |
| Activity requirements | 916 | 1,071 | 1,030 | 390 |
| No random coefficients | | | | |
| Reform | 797 | 1,328 | 846 | -34 |
| Increased generosity | -1,034 | -1,034 | -1,117 | -797 |
| Age limit | 207 | 314 | 204 | 74 |
| Activity requirements | 1,431 | 1,945 | 1,537 | 459 |
| Random coefficients: time, income, welfare | | | | |
| Reform | 416 | 704 | 482 | -145 |
| Increased generosity | -833 | -894 | -853 | -698 |
| Age limit | 119 | 30 | 165 | 104 |
| Activity requirements | 915 | 1,079 | 1,034 | 365 |

Notes: simulated effects of the reform when using different econometric specifications of the structural model.

Appendix

Table A1: Earnings Equation

| | Beta | St err |
|---------------------|-----------|--------|
| Hourly wage | | |
| Tertiary education | 7.066*** | 0.908 |
| Secondary education | 2.605*** | 0.655 |
| Lower education | | |
| Work experience | 0.431*** | 0.112 |
| Work experience sq. | -0.007*** | 0.002 |
| Part time job | -0.021 | 0.255 |
| Constant | 6.700*** | 2.164 |
| Selection | | |
| Tertiary education | 1.163*** | 0.120 |
| Secondary education | 0.518*** | 0.108 |
| Lower education | | |
| Work experience | 0.162*** | 0.011 |
| Work experience sq. | -0.003*** | 0.000 |
| Married/cohabitant | 0.260*** | 0.084 |
| Dependent children | -0.182*** | 0.070 |
| Household income | -0.026*** | 0.006 |
| Living in a city | 0.000 | 0.058 |
| Constant | -1.508*** | 0.146 |
| Lambda | 1.051 | 1.462 |
| Rho | 0.21 | |
| Observations | 2,667 | |

Notes: Heckman regression; *** significant at 1% level, ** significant at 5% level, *significant at 10% level. Hourly wage is expressed in €- 1998. Source: EU-SILC (2004).

Table A2: Other Benefits

| | Beta | St err |
|-----------------|----------|--------|
| One child | | |
| Two children | 3,076*** | 109 |
| Three children | 4,753*** | 141 |
| Four children | 5,818*** | 262 |
| Five children | 6,298*** | 518 |
| Six children | 6,807*** | 1017 |
| No work | 1,235*** | 271 |
| Short part time | 795*** | 238 |
| Part time | 695*** | 170 |
| Full time | | |
| Earnings / 1000 | -44*** | 10 |
| Immigrant | 1,147** | 465 |
| Constant | 6,635*** | 265 |
| Observations | 7,921 | |

Notes: Linear regression; *** significant at 1% level, ** significant at 5% level, *significant at 10% level.
Sample: lone mothers before the reform, 1995-1997.