Child Support Payments and Non-Compliance Cost: Does It Matter whether Money Comes from the Wallet or from the Purse?

Giedrius Blazys*

May 30, 2008

Abstract

Using Panel Study of Income Dynamics (1985-2005) data I construct a sample of divorced fathers who formed new partnerships. I use child support payment information to test the so-called "Income pooling" hypothesis, which is implied by Unitary Household Decision model. Under Unitary model household expenditures on the husband's children from his previous marriage should not be affected by intra-household income distribution. However, the new partner will likely receive less Utility from such expenditures, so her and his income will have different effect on child support payments if partners' relative incomes affect their bargaining power.

Although there is a great variation in fathers' payment behavior over years, and a large fraction of fathers don't pay any child support, a significant proportion of fathers pay what is ordered by court. Therefore, I jointly model father's decision to comply with child support orders and voluntary payment amounts to account for fathers who are simply paying what is ordered by court. Our estimates indicate that higher share of father's income in household income increases child support payment amounts. This finding rejects income pooling and is consistent with Family Bargaining models. However, the differential effect of father's income declines when controlling for individual heterogeneity in Random Effects regression, and it completely disappears in Fixed Effects Specification. Alternative explanations are suggested.

*Department of Economics, University of Washington, E-mail: giedrius@u.washington.edu. I am grateful to Shelly Lundberg, John Ermisch and to seminar participants at the University of Washington for their comments and suggestions. I also thank John Ermisch for suggesting the use of PSID dataset for the child support payments research topic, and Anne Case and Irwin Garfinkel for sharing State level Child Support policy variables.

1 Introduction

The second half of the previous century saw dramatic changes in how we understand a family, with increasing number of children raised by a single parent. Unprecedently high divorce rates in the last three decades and relatively low remarriage rates led to dramatic increase in the fraction of families with dependent children headed by single mothers. In the event of marital disruption, children traditionally stay with the mother and the father is expected to contribute to child rearing cost by paying child support.[1] Many fathers find new life partner and eventually get married again. It is intutitive to expect that, on average, higher father's income should lead to higher child support payments. However, what should be the relationship between father's child support payments and income from other family members in the father's household is less clear.

Traditional economic theory, which treats family as a single agent with explicit preferences and a single budget constraint, predicts that income source should not affect intra-household resource allocation, i.e. income from any family member should be spend in the same way as income from the other members. However, father's new partner most likely receives less utility from expenditures on his children from his previous marriage. If, instead, household resources are allocated as described by Cooperative Bargaining models and if partners' relative income affect their bargaining power, then child support payments will be affected by the variation in the source of household income.

A vast body of the studies about child support payment behavior in the United States use data from Current Population Survey (CPS) or The Survey of Income and Program Participation (SIPP). These large nationally representative datasets provide detailed information about custodial mothers' characteristics and their reports about child support award and payments amounts. Unfortunately, these datasets contains virtually no information about noncustodial fathers' income and other characteristics and even less so about fathers' new partners. However, Smock and Manning (1997) argue that nonresident parent's characteris-

[1] A mother becomes a Custodial parent in about 90% of divorce settlement cases

tics are more important when describing child support payment behavior than the resident parent's characteristics. They use matched resident and nonresident parent data and find that including resident parent's characteristics add very little to the predictive power of child support payment regressions. Thus, having information on the nonresident parents is essential if we want to analyze child support payments and compliance with court orders. In this paper I use Panel Study of Income Dynamics (PSID) data. To my knowledge, PSID is the only large representative US dataset which contains information about child support payments and nonresident fathers' characteristics as well as characteristics of the other household members.

Income pooling hypothesis has been tested and in general rejected in a variety of settings.[2] This study is largely motivated by Ermisch and Pronzato (2008). Ermisch and Pronzato use British panel data to construct a sample of divorced or separated fathers with dependent children. They consider fathers who are remarried or in cohabiting relationship with another women and find that higher share of father's income in household income increases both the probability of child support payment and child support share relative to the household income, and thus they are able rejects income pooling.

However, child support payments can only have a behavioral interpretation if they are made voluntarily.[3] Ermisch and Pronzato argue that high prevalence of informal child support arrangements and weak child support order enforcement in the United Kingdom allows them to assume that child support payments are voluntary. National statistics suggest that situation in the United States is not much different - Census staff estimates that in the U.S. only about 60% of previously married mothers have formal child support awards, of which only about 45% receive the full amount awarded. This low level of compliance might be a good indicator that child support transfers are to large degree voluntary.

[2]Researchers, for example, used individual data on leisure times or labor supplies, and even expenditures on men and women's clothing and tested if they depend on the variation in the source of household income. See, e.g. Fortin and Lacroix 1997; Lundberg et al. 1997; Chiappori et al. 2002.

[3]If fathers are just paying the amount specified by the child support court order, child support essentially becomes just another income tax.

Nevertheless, there still is a significant fraction of fathers who are actually paying what is ordered by court. Moreover, the United States government increased its effort to collect child support orders in the late 80's and 90's.[4] Therefore, I jointly model voluntary support payment amounts and father's decision to comply (or not to comply) with child support order, in order to account for the fact that some fathers are simply paying what is ordered by court.[5] Arguably, noncutodial fathers incur some monetary or nonmonetary cost if they decide to pay less than what is ordered by court (including zero payments). If these noncompliance costs are greater than the loss in utility resulting from paying what is ordered by court, fathers decide to comply with the court order and for such fathers voluntary child support payments are not ubserved. On the other hand, if the father is paying significantly more or significantly less than the court order amount, I assume that such payments are voluntary. The next section of the paper presents a theoretical model which provides motivation for my empirical analysis.

2 The Theoretical Model

High divorce rates result in a large proportion of children living with one of their biological parents, while the other parent is expected to contribute to child rearing expenses by providing some form of monetary or nonmonetary support. Since the mother is the custodial parent in a large majority of child support cases, I use a term "mother" interchangeably with "custodial parent" and "father" with "non-custodial parent". I assume that both divorced parents care about their children welfare, so child quality remains, in a sense, a public good after parents get divorced (see Weiss and Willis (1985, 1989) for further discussion). Following Del Boca and Flinn (1995), I assume that the mother is solely responsible for making expenditures on children, while fathers can raise children consumption only indirectly

[4]See Garfinkel et al. 2001 for a list of papers providing empirical evidence about the effects of child support enforcement efforts on child support payments and compliance levels.

[5]In a sense, I model such fathers using selection framework. My econometric specification also allows for the fact that about 50% of noncustodial fathers do not pay any child support at all.

through money transfers to the mother.

The main difference between my paper and Del Boca and Flinn (1995), is that I consider fathers who formed new families by either re-marrying or entering a cohabitation relationship, while Del Boca and Flinn (1995) assume that fathers remain single. Let c_m, c_f and c_p denote private consumption levels of the mother, the father and his new life partner, correspondingly. Also let k stand for child good expenditures and t be father's transfer amount to mother's household. I assume that individual Utility functions take a simple Cobb-Douglas functional form. The mother maximizes her Utility function subject to the budget constraint:

$$\max_{c_m,k} U_m = \delta_m \log(c_m) + (1-\delta_m)\log(k), \quad s.t. \ c_m + k = y_m + t, \tag{1}$$

where δ_i is the preference parameter towards private consumption of parent $i = m, f$. This results in her optimal consumption level of $c_m^* = \delta_m (y_m + t)$ and child expenditures $k^* = (1-\delta_m)(y_m + t)$.

I assume that the father and his new partner decide how to allocate their resources through a bargaining process - a modeling approach pioneered by Manser and Brown (1980) and McElroy and Horney (1981). If reaching the agreement between spouses is not too costly, family members can potentially achieve the cooperative equilibrium. In what follows, I do not consider how the bargaining process takes place and therefore I follow the "Collective" modeling approach as suggested by Apps and Rees (1997) and Chiappori (1988). By assumption, this "collective" equilibrium is Pareto efficient, i.e. we cannot improve one spouse's situation without hurting the other spouse[6]. However, the actual outcome in this collective equilibrium depends on the bargaining power of each spouse, which is captured by parameter μ.

Moreover, I assume that father is expected to pay at least s amount of child support which is stipulated by the court order, and failure to do so results in fixed noncompliance

[6]Chiappori (1992) was among the first to utilize the fact that finding Pareto efficient intrahousehold allocations is equivalent to maximization of weighted sum of individual utilities, where μ can be interpreted as Lagrange multiplier associated with the Pareto efficiency constraint: $\max U^1(x_1)$, s.t. $U^2(x_2) \geq \bar{u_2}$.

cost of ϑ. These cost are expressed in terms of utils and they could be monetary (such as the penalty if father is "caught" noncomplying) or nonmonetary (such as guilt, reduced time spent with the child, etc.). I assume that father's new partner does not derive any utility from expenditures on the man's children from his previous marriage. The cooperative solution can be found by maximizing family's welfare function, which is formulated as a weighted sum of individual spouses' utilities, subject to a pooled income budget constraint, and mother's expenditures on child quality:

$$\begin{aligned} \max_{c_f, c_p, t} U_f + \mu U_p &= \delta_f \log(c_f) + (1-\delta_f)\log(k) - \vartheta I[t < s] + \mu \log(c_p), \\ s.t. \quad & y_f + y_p = t + c_f + c_p, \\ & k = (1-\delta_m)(y_m + t). \end{aligned} \tag{2}$$

where $I\,[]$ is an indicator function, which shows that father's household incurs noncompliance cost ϑ only if father decides to pay less child support than what was ordered by the court.

The solution of the father's household utility maximization problem is provided in the Appendix A. Optimal voluntary child support transfer value is given by:

$$t^* = \frac{1-\delta_f}{1+\mu}(y_f + y_p) - \frac{\mu + \delta_f}{1+\mu} y_m \tag{3}$$

As indicated by equation (3), voluntary child support payment depends on joint father's and his partner's income and Pareto weight μ that measures the bargaining power of each spouse. In Unitary household decision models this Pareto weight is fixed, while Collective models suggest that it should depend on prices, individual income, and other so-called "distribution factors". Therefore, testing if the effects of father's and his partner's income on child support payments are different is equivalent to testing the Unitary model versus a more flexible Collective modeling approach. This test is generally referred to as the test of "Income Pooling" hypothesis.

Depending on the values of father's preference and noncompliance cost parameters, father may decide to pay no child support, which I call a “No Payments” case, he might be willing to pay less than the court order amount, which I refer to as a “Partial Payments” case, he might decide to pay exactly what is ordered by court, which I call an “Exact Compliance” case, and finally, he might be willing to pay more than what is ordered by court, which I refer to as an “Over Compliance” case. The solution for all these cases is provided in Appendix A and can be summarized by the following equation system:

$$
\begin{array}{llll}
1)\ \textit{No Payments} & t=0 & if & \delta_f \in \left(\overline{\delta},1\right] \ and\ \vartheta \in \left[0, W^{NP}-W^{EC}\right), \\
2)\ \textit{Partial Payments} & t=t^*<s & if & \delta_f \in \left(\underline{\delta},\overline{\delta}\right] \ and\ \vartheta \in \left[0, W^{PP}-W^{EC}\right), \\
3)\ \textit{Exact Compliance} & t=s & if & \begin{array}{l} \delta_f \in \left(\overline{\delta},1\right] \ and\ \vartheta \in \left[W^{NP}-W^{EC},\infty\right) \\ \delta_f \in \left(\underline{\delta},\overline{\delta}\right] \ and\ \vartheta \in \left[W^{EC},\infty\right) \end{array}, \\
4)\ \textit{Over Compliance} & t=t^*>s & if & \delta_f \in \left[0,\underline{\delta}\right],
\end{array}
\tag{4}
$$

where $\overline{\delta} \equiv 1-(1+\mu)\frac{y_m}{y_T}$ and $\underline{\delta} \equiv 1-(1+\mu)\frac{(y_m+s)}{y_T}$ are the threshold values for father's preference parameter, while W^{EC}, W^{PP} and W^{NP} denote indirect utility values (without noncompliance cost) in “Exact Compliance”, “Partial Payments” and “No Payments” case (the actual expressions are provided in Appendix A).

As equation system (4) indicates, we can only observe voluntary child support payment behavior when fathers pay less (including zero payment) or more child support than the court order amount. When father is complying with the court order, $t = s$, voluntary child support payment amount, t^*, is not observed, and we can only infer that it is less than or equal to the order amount: $t^* \leq s$. Therefore, if we want to test the “Income Pooling” hypothesis using child support payments, we will have to account for “selection” of fathers into “Exact Compliance” regime. Moreover, in the empirical part of the paper I also have to control for the fact that a large proportion of fathers choose not to pay any child support. In the next section of this paper I propose an econometric specification, which models fathers' “selection” into “Exact Compliance” and allows for zero child support payments.

It should be noted, that although I refer to ϑ as the “fixed” noncompliance cost, it is not the same for different fathers and it does not have to be constant over time. Following Del Boca and Flinn (1995), by "fixed" I assume that these costs do not depend on the compliance level, $s - t^*$. We can think of these costs, as the costs of breaking a promise or obligation to pay a certain amount of child support, no matter what is the size of the obligation, or what is the size of arrears.

3 The Econometric Model

3.1 Individual Heterogeneity Specified as Random Effects

Let y_{1it}^* be the unobserved, or latent, voluntary child support payment amount, while y_{2it}^* denotes latent variable which measures noncompliance costs (where higher values of y_{2it}^* indicate lower costs). Also let s_i be individual (predetermined) child support court order amount. Consider the following model:

$$\begin{cases} y_{1it}^* = \beta' x_{it} + \sigma_\epsilon \epsilon_i + \nu_{it} \\ y_{2it}^* = \gamma' z_{it} + \sigma_u u_i + \omega_{it} \end{cases}, \tag{5}$$

where $\begin{pmatrix} \epsilon \\ u \end{pmatrix} = N\left(0, \begin{pmatrix} 1 & \\ \rho_h & 1 \end{pmatrix}\right)$ are correlated individual heterogeneity terms and $\begin{pmatrix} \omega \\ \nu \end{pmatrix} = N\left(0, \begin{pmatrix} 1 & \\ \rho\sigma_\nu & \sigma_\nu^2 \end{pmatrix}\right)$ are correlated contemporaneous errors. Both error components are assumed to be uncorrelated with observable independent variables.

We observe actual voluntary child support payment amount, $y_{it} = y_{1it}^*$, if it is higher than the order amount ($y_{1it}^* > s_i$) or if its lower than the order amount, and father does not comply ($y_{1it}^* < s_i, y_{2it}^* > 0$). If observed child support payment is equal to the order amount ($y_{it} = s_i$), than we know that voluntary child support payment is less or equal to the order amount ($y_{1it}^* \leq s_i$) and that non-compliance cost are prohibitively high ($y_{2it}^* \leq 0$), so fathers are paying what is ordered by the court. If voluntary support payment is higher than the order amount, we do not know anything about compliance cost, since for such father compliance issue is irrelevant.

This model is estimated using Maximum Likelihood Estimator (MLE):

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \log L\left(\theta | \, data\right), \tag{6}$$

where θ is the full parameter vector $\theta = [\beta; \gamma; \rho_h; \sigma_\epsilon; \sigma_u; \rho; \sigma_\nu]$, observed $data = [Y; X; Z]$, while $L\left(\theta | \, data\right)$ is the likelihood function for the sample.

The probability density function for each observation can be decomposed into four different parts, depending on the observed child support payments:

$$y_{it} = \begin{cases} 0 & \text{if } y^*_{1it} \leq 0 \text{ and } y^*_{2it} > 0; \\ y^*_{1it} & \text{if } y^*_{1it} > 0 \text{ and } y^*_{1it} < s_i \text{ and } y^*_{2it} > 0; \\ s_i & \text{if } y^*_{1it} \leq s_i \text{ and } y^*_{2it} \leq 0; \\ y^*_{1it} & \text{if } y^*_{1it} > s_i. \end{cases} \tag{7}$$

Density function for each of these four cases is derived in Appendix B. The density function for any y_{it}, conditional on individual heterogeneity effects, is the product of the densities for these 4 parts weighted by indicator functions:

$$\begin{aligned} f\left(y_{it} | \, u_i, \epsilon_i\right) &= \left\{ \Phi_2 \left(\tfrac{-\beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu}, \gamma' z_{it} + \sigma_u u_i, -\rho \right) \right\}^{I(y_{it}=0)} \\ &\times \left\{ \Phi \left(\tfrac{\gamma' z_{it} + \sigma_u u_i + \frac{\rho}{\sigma_\nu}(y_{it} - \beta' x_{it} - \sigma_\epsilon \epsilon_i)}{(1-\rho^2)^{1/2}} \right) \tfrac{1}{\sigma_\nu} \phi \left(\tfrac{y_{it} - \beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu} \right) \right\}^{I(0 < y_{it} < s_i)} \\ &\times \left\{ \Phi_2 \left(\tfrac{s_i - \beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu}, -\gamma' z_{it} - \sigma_u u_i, \rho \right) \right\}^{I(y_{it}=s_i)} \\ &\times \left\{ \tfrac{1}{\sigma_\nu} \phi \left(\tfrac{y_{it} - \beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu} \right) \right\}^{I(y_{it} > s_i)} \end{aligned} \tag{8}$$

To get the unconditional densities, we need to "integrate out" individual heterogeneity terms ϵ_i and u_i. Since conditioned on ϵ_i and u_i, the $y_{it}s$ are assumed to be independent, we have

$$f\left(y_{i1}, y_{i2}, \ldots | \, u_i, \epsilon_i\right) = f\left(\mathbf{Y}_i | \, u_i, \epsilon_i\right) = \prod_t f\left(y_{it} | \, u_i, \epsilon_i\right) \tag{9}$$

Then the unconditional distribution is

$$
\begin{aligned}
f(\mathbf{Y}_i) &= \mathbf{E}_{u,\epsilon}\left[f\left(\mathbf{Y}_i|\, u_i,\epsilon_i\right)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\left(\mathbf{Y}_i|\, u_i,\epsilon_i\right) g\left(u_i,\epsilon_i\right) du_i d\epsilon_i \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \prod_t f\left(y_{it}|\, u_i,\epsilon_i\right) g\left(u_i,\epsilon_i\right) du_i d\epsilon_i,
\end{aligned}
\tag{10}
$$

where $g(u_i,\epsilon_i)$ is a joint pdf, which is assumed to be bivariate normal. This expectation is estimated using Gauss-Hermite quadrature, which essentially "discretizes" this joint pdf by replacing it with joint distribution of discrete random variables with mass points (or nodes of approximation) u_m and ϵ_l and probability weights W_{ml}:

$$
f(\mathbf{Y}_i) \approx \sum_m \sum_l W_{ml} \prod_t f\left(y_{it}|\, u_m,\epsilon_l\right) \tag{11}
$$

In order to assure that σ_ν^2, σ_ϵ^2, and σ_u^2 are positive, for computational reasons, they are reparameterized as $\sigma_j^2 = \exp(\alpha_j)$, where $j = \nu, \epsilon, u$. In addition, in order to impose the restriction $(-1 < \rho < 1)$, I reparameterize $\rho = \frac{1-\exp(\alpha_\rho)}{1+\exp(\alpha_\rho)}$. Then $\frac{1}{(1-\rho^2)^{1/2}} = \frac{1+\exp(\alpha_\rho)}{2\exp\left(\frac{1}{2}\alpha_\rho\right)}$. Similarly, ρ_h is reparameterized as $\rho_h = \frac{1-\exp(\alpha_h)}{1+\exp(\alpha_h)}$. Let $\tilde{\theta}$ denote the parameter vector without individual heterogeneity correlation parameter:

$$
\tilde{\theta} = \left[\beta; \gamma; \alpha_\rho; \alpha_\nu; \alpha_\epsilon; \alpha_u\right].
$$

Similarly as in Greene (1998), for computational reasons $f(\mathbf{Y}_i)$ is estimated as:

$$
f(\mathbf{Y}_i) \approx \sum_m \sum_l W_{ml} \exp\left(\sum_t l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)\right), \tag{12}
$$

where $l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right) \equiv \log\left(f\left(y_{it}|\, u_m,\epsilon_l\right)\right)$.

Then log-likelihood for the whole sample is

$$
l(\theta) = \sum_i l_i(\theta) \approx \sum_i \log\left(\sum_m \sum_l W_{ml} \exp\left(\sum_t l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)\right)\right), \tag{13}
$$

where $l_i(\theta) \equiv \log(f(\mathbf{Y}_i))$.

The gradient of the sample log-likelihood function is estimated as

$$\frac{\partial l(\theta)}{\partial \theta} = \sum_i \frac{\frac{\partial f(\mathbf{Y}_i)}{\partial \theta}}{f(\mathbf{Y}_i)}, \tag{14}$$

where $\partial f(\mathbf{Y}_i)/\partial \tilde{\theta}$ is approximated by the following expression:

$$\frac{\partial f(\mathbf{Y}_i)}{\partial \tilde{\theta}} \approx \sum_m \sum_l W_{ml} \exp\left(\sum_t l_{it}\left(\tilde{\theta}\,|u_m, \epsilon_l\right)\right)\left(\sum_t \frac{\partial l_{it}\left(\tilde{\theta}\,|u_m, \epsilon_l\right)}{\partial \tilde{\theta}}\right) \tag{15}$$

While expression for $\partial f(\mathbf{Y}_i)/\partial \alpha_h$ is given by:

$$\frac{\partial f(\mathbf{Y}_i)}{\partial \alpha_h} \approx \sum_m \sum_l C_{ml}^h W_{ml} \exp\left(\sum_t l_{it}\left(\tilde{\theta}\,|u_m, \epsilon_l\right)\right), \tag{16}$$

where $C_{ml}^h \equiv -\frac{\rho_h}{2}\left(1 - \frac{u_m^2 + \epsilon_l^2 - 2\rho_h u_m \epsilon_l}{\left(1-\rho_h^2\right)} + \frac{u_m \epsilon_l}{\rho_h}\right)$

Maximum Likelihood parameters are found by setting the gradients equal to $\mathbf{0}$, i.e. by solving the likelihood equation:

$$\frac{\partial \log L}{\partial \theta} = \mathbf{0} \tag{17}$$

The asymptotic covariance matrix of the estimated coefficients is computed using the BHHH estimator, which is estimated as the sum of the outer products of the gradients. The reparameterized coefficients and their standard errors are estimated using the Delta method.

3.2 Fixed Effects estimation

Alternatively, we can solve the model specified in equation (5), where σ_ϵ and σ_u is set to 1, using Fixed Effects (FE) estimation. When time dimension is fixed, parameter estimates will not be consistent even in large panel data samples, since as the number of observations increases, so does the number of parameters to be estimated (incidental variables problem).

However, when T is sufficiently large, the bias might be small and of little practical importance, especially if regressions do not include lagged dependent variables (see Heckman 1981 and Honore 1993 for evidence using Monte Carlo simulations for Fixed Effects Probit and Tobit models). The main advantage of the FE is that we do not have to assume that unobserved individual heterogeneity is uncorrelated with our regressors, which is a maintained assumption when estimating the model using Random Effects specification as in the previous section.

When individual heterogeneity terms are estimated as Fixed Effect, the sample log-likelihood of the model becomes the following:

$$l\left(\theta, \eta_1, \ldots, \eta_N\right) = \sum_i l_i\left(\theta, \eta_i\right) = \sum_i \sum_t l_{it}\left(\theta, \eta_i\right), \tag{18}$$

where $\theta = [\beta; \gamma; \sigma_\nu]$, $\eta_i = [\epsilon_i; u_i]$, while $l_{it}\left(\theta, \eta_i\right)$ is as defined in equation (36) in Appendix B.

We can estimate this model by maximizing concentrated log-likelihood function, where individual fixed effects, η_i, are "concentrated out" of the log-likelihood. This is accomplished by finding the MLE of η_i for given θ:

$$\hat{\eta_i}\left(\theta\right) = \arg\max_{\eta_i} l_i\left(\theta, \eta_i\right), \tag{19}$$

and then substituting $\hat{\eta_i}\left(\theta\right)$ into the sample log-likelihood and maximizing it with respect to θ:

$$\hat{\theta} = \arg\max_{\theta} \sum_i l_i\left(\theta, \hat{\eta_i}\left(\theta\right)\right) \tag{20}$$

This two step estimation procedure is iterated by re-estimating η_i for given $\hat{\theta}$ and then estimating θ for new $\hat{\eta_i}$. This iteration is continued until the change in $\hat{\theta}$ is smaller than some specified criterion.

Denote score functions $d_{\eta i}\left(\theta, \eta_i\right) \equiv \frac{\partial l_i(\theta,\eta_i)}{\partial \eta_i}$ and $d_{\theta i}\left(\theta, \eta_i\right) \equiv \frac{\partial l_i(\theta,\eta_i)}{\partial \theta}$. Expressions for these derivatives are defined similarly as for $\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)/\partial\tilde{\theta}$, which are given in the Appendix B. Then $\hat{\eta}_i\left(\theta\right)$ is estimated by solving

$$d_{\eta i}\left(\theta,\eta_i\right)=\sum_t \frac{\partial l_{it}\left(\theta,\eta_i\right)}{\partial\eta_i}=\mathbf{0}, for\ \ i=1,\ldots,N \tag{21}$$

while $\hat{\theta}$ solves the following first order conditions

$$\sum_i\left[d_{\theta i}\left(\theta,\hat{\eta}_i\left(\theta\right)\right)+\frac{\partial\hat{\eta}_i\left(\theta\right)'}{\partial\theta}d_{\eta i}\left(\theta,\hat{\eta}_i\left(\theta\right)\right)\right]=\sum_i d_{\theta i}\left(\theta,\hat{\eta}_i\left(\theta\right)\right)=\mathbf{0} \tag{22}$$

Variance-Covariance matrix of $\hat{\theta}$ is estimated as minus the inverse of the Hessian matrix. Following Carro (2007), the Hessian of the concentrated log-likelihood is adjusted for the fact that individual fixed effects are estimated (See Appendix C for derivation):

$$\frac{\partial^2 l^C\left(\theta\right)}{\partial\theta\partial\theta'}=\sum_i\left[d_{\theta\theta i}\left(\theta,\hat{\eta}_i\left(\theta\right)\right)-d_{\theta\eta i}\left(\theta,\hat{\eta}_i\left(\theta\right)\right)\left[d_{\eta\eta i}\left(\theta,\hat{\eta}_i\left(\theta\right)\right)\right]^{-1}d_{\theta\eta i}\left(\theta,\hat{\eta}_i\left(\theta\right)\right)'\right], \tag{23}$$

where $d_{\theta\theta i}\left(\theta,\eta_i\right)\equiv\sum_t\frac{\partial^2 l_{it}\left(\theta,\eta_i\right)}{\partial\theta\partial\theta'}$, $d_{\theta\eta i}\left(\theta,\eta_i\right)\equiv\sum_t\frac{\partial^2 l_{it}\left(\theta,\eta_i\right)}{\partial\theta\partial\eta_i'}$ and $d_{\eta\eta i}\left(\theta,\eta_i\right)\equiv\sum_t\frac{\partial^2 l_{it}\left(\theta,\eta_i\right)}{\partial\eta_i\partial\eta_i'}$ are estimated by numerically differentiating score functions. The adjustment factor (the second term in equation 23) is set to 0 for a few cases when individual effect in noncompliance equation (the second element of η_i) is not identified. This happens when fathers always over-comply.

4 Data and Sample Construction

I estimate the model using data from the Panel Study of Income Dynamics (PSID) survey.[7] PSID is a longitudinal survey of a representative sample of US households started in 1968 and is still ongoing. I use both the nationally representative sample and the low-income

[7]The Panel Study of Income Dynamics is primarily sponsored by the National Science Foundation, the National Institute of Aging, and the National Institute of Child Health and Human Development and is conducted by the University of Michigan.

families sample in the analysis. One of the main advantages of this data for the purpose of this study is the availability of detailed childbirth and marriage histories as well as a plethora of yearly socio-economic indicators. Information about household head's and wife's annual child support payments is available since 1985 and refers to the previous calendar year[8]. I restrict the sample to include only fathers who are household heads and who are living together with another woman who is not a child's mother.

Noncustodial parents are obliged to provide child support until age 18 or 19 (a few states have a termination clause upon emancipation of the minor) (National Conference of State Legislatures 2007). I use childbirth and marriage histories to determine biological children from previous marriages, who are below age 18, and who are living outside the father's household and thus are "at risk" of receiving child support. Coresidence information is reported at the time of the interview, while most socio-economic variables, including income and child support payments, refer to the previous calendar year. Matching coresidence information to the income and support payment information of the same calendar year is problematic, since starting from 1997, PSID became biennial. Moreover, any changes in family structure and coresidence status recorded at the time of the interview could have happened at any time in the preceding year (Page and Stevens 2004). Therefore, I ignore the different time reference and match family coresidence status as well as child support and income data from the same survey year.[9]

PSID contains data on child support payments, but no information about child support court orders. Historically, child support orders were based on the perceived child's needs and father's ability to pay, and were set on a case by case basis. This often resulted in

[8]PSID defines both head's legal wife and his cohabiting partner as head's wife and collects information about her. Since PSID dataset defines cohabitation as a "long-term" relationship, they do not include first year partners as cohabitors. Thus, I redefine individuals as cohabitors if they are present in the household in the current year and are defined as cohabiting in the next year.

[9]When I match current survey year's coresidence status information and next year's lagged income information, in many cases, I lose one year of observations per individual, which results in smaller sample sizes, especially in Fixed Effects estimation. Regression analysis when using this matching approach leads to qualitatively similar coefficients but larger standard errors. Most of the Fixed Effects coefficients become statistically insignificant.

relative inconsistency among cases and somewhat low orders (National Women's Law Center 2002). The Child Support Enforcement amendments of 1984 required each state to develop a numeric guideline which could be used to calculate child support orders. The Family Support Act of 1988 required the States to start using these guidelines universally, except for special cases. Moreover, all States were required to enact statutes providing for the use of improved enforcement mechanisms, like mandatory income withholding or State income tax refund interceptions (National Women's Law Center 2002).

Since starting from the late 80's courts have to use guidelines to set order amounts, I use State specific guideline amounts as a proxy for child support orders. I am predicting guideline amounts using data from Pirog et al. (1998) paper which contains State level child support guideline amounts for hypothetical income scenarios for years 1991, 1993, 1995 and 1997, and similar data from Morgan and Lino (1999) paper for the year 1999. I interpolate and extrapolate the implied guideline schedules for the remaining years in my data sample.[10]. I use average father's income around the time of divorce or separation and information about father's state and number of nonresident dependent children in each survey year to predict child support court order. This is equivalent to assuming that order amount is adjusted to reflect the changes in the number of nonresident dependent children and state guideline schedules but is not adjusted for changes in father's income.[11] Moreover, guidelines specify adjustments to account for shared-parenting time, child care and medical expenses. However, such information for nonresident children is not available in the PSID dataset, so my predicted court order will involve a significant measurement error.[12]

[10]Although federal laws require states to review and, if found appropriate, to modify guideline formulas at least once every 4 years, such modifications are not done frequently and some states have not updated their guideline formulas for years (Venohr and Griffith 2005)

[11]When there is a significant change in financial situation of custodial or noncustodial parent, either of them can request child support award to be modified; however, such modifications or eliminations of awards are rare (Peterson and Nord 1990). Although, OCSE periodically reviews child support orders for mother who receive welfare payments, my sample consist mostly of non-welfare cases. Only 5% of mothers in the matched mothers and fathers sample report receiving AFDC (or TANF after 1996) income.

[12]A significant portion of mothers choose not to obtain a child support award. Since I do not have information if the mother actually was granted a child support award, for such cases the predicted guideline amount will measure the potential court order amount and not the actual award. These cases are still consisted with our model where we allow a positive court order to exist and implicit noncompliance cost to

Table 1: Compliance with child support orders

No payments	48%
Partial Payments	32%
Full Compliance	14%
Over Compliance	6%

Sample size: 3414 person-years, 957 individuals. Observations are weighted using PSID Household weights. Source: Author's estimation using 1985-2005 PSID data

To identify cases where fathers just comply with child support orders, as opposed to choosing payment amount voluntarily, I compare the actual payment amount with the order amount. I assume that fathers are fully complying if their child support payments are close to the predicted court order amount (within 20% of the order). As table 1 shows, about 14% of fathers are identified as fully complying with the court order and 6% are predicted to be voluntarily paying more than ordered by court. The percentage of fathers who are paying at least what is ordered by court (14+6=20%) is somewhat lower than suggested by nationally representative CPS data, since as table 2 indicates, about 30% of mothers are receiving full payments.[13] This again suggest that I might be misclassifying some cases when fathers pay voluntarily vs. simply complying with court orders. I discuss the implications of such misclasification in the last section of the paper.

In all regressions I use total money income, which is the sum of labor, asset, and transfer income. Although in some cases it might be difficult to assign asset income to a specific individual in the household[14], the use of just labor income might not be satisfactory, since child support orders are based on father's total income.[15] Extreme child support and income

be zero.

[13]Note that by survey design "Full compliance" category in CPS includes fathers who actually "over-comply", so we cannot use CPS data to infer what percentage of fathers are paying more than what is ordered by court. Smaller scale dataset containing actual order and payment information suggest that a considerable fraction of fathers do. For example, Del Boca and Flinn (1995) using Court record data from Wisconsin for years 1980-1982 estimated that five months after the divorce decree 40% of Noncustodial fathers could classified as exact compliers and 11% paid significantly more than ordered by the court.

[14]If the owner of an asset, which is the source of income, is not reported, I divide such income to head and wife of the household equally.

[15]Some states use Gross total income, while other states use after-tax total income to estimate basic child

Table 2: Ever married Custodial Mothers by Child Support Receipts Status in 2002 April CPS

	Number in 1000's	Percent
Ever married Custodial mothers	7768	100%
With child support agreements or awards	5276	68%
Due child support payments in 2003	4640	60%
Received full payments	2313	30%
Received part payments	1316	17%
No payments or no award	4139	53%

source: Author's estimation using data from Table 8 in Grall (2006)

values are censored to lower the impact of possible measurement error, and I drop observations where total reported father's household income is lower than $100 a month. After excluding observation with missing information on main characteristics like income, child support payment amount, or race, we are left with 957 individuals with at most 17 years of data resulting in 3414 person-years observations.

Table 3 shows descriptive statistics of the main sample. Throughout the paper, all monetary amounts are expressed in terms of 2000 dollars. Including zero payments, average child support payments are less than $3000, which is about 4.3% of total household income. More than half of all fathers pay some child support, so paying fathers, on average, transfer almost $5600 per year. More than 40% of fathers have at least four-year college degree and about 75% of them are married to their new life partner. Other variables used in empirical analysis include the number of father's biological children and the number of other children in the new household, years since divore or separation, and a dummy for more than one dependent children living outside father's household.[16]

I assume that the following variables affect non-compliance cost, but not the actual

support orders.

[16]As implied by the theoretical model, father's child support payments should be affected by mother's (custodial parent's) income and her other characteristics. Unfortunately, mother's time invariant characteristics are available for less than half of father's observations in my sample, and time varying - only for 20% of observations. PSID only follows the so called *sample* individuals, who are from the original 1968 sample or are offspring of the original sample members. Thus, in most cases, only a father or only a mother can be a *sample* member in later survey years. So, in the sample of nonresident fathers, mothers' information, in general, will not be available.

Table 3: Descriptive Statistics

Variable	Mean	Std.Dev.
Child Support Payment Amount (1,000's)	2.90	4.04
Order amount (1,000's)	7.25	4.98
Total Income (10,000's)	6.75	4.05
Father's Income (10,000's)	4.37	2.98
# of own children in the HH	0.57	0.85
# of other children in the HH	0.58	0.89
Years since marriage ended	8.29	4.00
OCSE expenditures per single mother (1,000's)	0.40	0.21
If father married again	0.75	
Proportion paying some child support	0.52	
If two or more dependent children outside the HH	0.37	
Proportion non-white	0.11	
Proportion with college degree	0.41	
Proportion self employed	0.15	
Proportion working in public sector	0.12	
If state has immediate income withholding	0.71	
If state adopted numerical guidelines	0.74	

Sample size: 3414 person-years, 957 individuals. Observations are weighted using PSID Household weights. Money amounts are in Constant 2000 dollars.

child support payment: if father is self employed; if father working is in public sector; and state level child support enforcement characteristics.[17] I use three variables to measure the strength of child support enforcement policy: a dummy indicating if a state has immediate income withholding from non-resident parents' earnings when these parents miss or are likely to miss payments; a dummy for presumptive guidelines – if states are required to use numeric guidelines for setting child support awards; and state expenditures on enforcement – the expenditures reported by OCSE divided by the number of single-mother families in a particular state (see Aizer and McLanahan 2006, Case et al. 2003, Garfinkel et al. 2001, Freeman and Waldfogel 2001, or Sorensen, Elaine and Halpern, Ariel 1999 for further discussion about the use of these variables).

[17] Thus these variables are among $Z's$, but excluded from $X's$. This exclusion restriction helps identification of regression parameters.

5 Results

Results from the model without individual heterogeneity (Pooled) and the model where individual heterogeneity is specified as random effects (RE) are shown in Table 4. Estimated variances of heterogeneity terms in both voluntary payments and noncompliance selection regressions are highly significant, which indicates the presence of heterogeneity effects. Coefficient estimates from both Pooled and RE regressions indicate significantly positive effects of total household income and individual father's income on voluntary child support payments, which suggests that increase in father's income has differential effect than increase in his partner's income, which rejects the "income pooling" hypothesis and is consistent with Ermisch and Pronzato (2008) results. However, the difference in the effect of father and his partner incomes on child support payment is smaller in Random Effects regression, and it completely goes away in the Fixed Effects Specification.

Results from the Fixed Effects regression are listed in table 5. To lower the possible inconsistency of the Maximum Likelihood estimates I restrict the sample to individuals who have at least 4 years of observations, which results, on average, in 7 observations per individual. As table 5 shows, when the model is estimated assuming FE specification, father's individual income effect is very small and statistically insignificant. This could be interpreted as a sign that if families behave as predicted by cooperative bargaining model, then yearly variation in income might not be a good indicator of differences in bargaining powers, i.e. "permanent" income component or potential income matters more. Alternatively, this could suggest a bias in Pooled and RE models due to unobserved heterogeneity, if for example, more productive fathers are also more responsible and care about their children. In this case FE will be unbiased and would imply that families actually pool their resources.

Table 4: **Full Sample Estimation Results**

Variables	**No Individual Heterogeneity**	**Random Effects**
Child Support Payment Amount (1,000's)		
Constant	-4.851**	-4.006**
	(0.354)	(0.532)
Total Income (10,000's)	0.238**	0.236**
	(0.057)	(0.079)
Father's Income (10,000's)	0.450**	0.275**
	(0.070)	(0.100)
# of own children in the HH	-0.472**	-0.318+
	(0.118)	(0.183)
# of other children in the HH	-0.390**	-0.413*
	(0.104)	(0.168)
If father married again	1.003**	0.974**
	(0.232)	(0.329)
Years since previous	-0.137**	-0.139**
marriage ended	(0.024)	(0.035)
If two or more dependent	2.142**	1.662**
children outside the HH	(0.204)	(0.302)
If father not white	-2.114**	-2.070**
	(0.229)	(0.419)
If father has college degree	1.431**	1.941**
	(0.191)	(0.378)
$y_2 = 1$ if does not comply with child support order		
Constant	1.162**	1.610**
	(0.116)	(0.205)
Total Income (10,000's)	-0.032*	-0.049+
	(0.016)	(0.026)
Father's Income (10,000's)	-0.038+	-0.023
	(0.019)	(0.032)
# of own children in the HH	0.119**	0.106
	(0.043)	(0.070)
# of other children in the HH	0.107**	0.123*
	(0.030)	(0.052)
If father married again	-0.011	-0.004
	(0.063)	(0.102)
Years since previous	0.052**	0.057**
marriage ended	(0.008)	(0.013)

Continued on next page

Table 4 – continued from previous page

Variables	No Individual Heterogeneity	Random Effects
If two or more dependent children outside the HH	-0.180**	-0.194*
	(0.055)	(0.089)
If father not white	0.290**	0.304*
	(0.080)	(0.137)
If father has college degree	0.010	-0.039
	(0.062)	(0.118)
OCSE expenditures per single mother	0.007	-0.193
	(0.135)	(0.211)
If state has immediate income withholding	0.224*	0.255+
	(0.102)	(0.151)
If state adopted numerical guidelines	-0.236*	-0.279+
	(0.118)	(0.160)
If father self employed	0.427**	0.434**
	(0.104)	(0.152)
If father works in public sector	-0.302**	-0.373**
	(0.077)	(0.133)
σ_v	5.919**	4.376**
	(1.478)	(0.646)
ρ	0.419+	0.432**
	(0.253)	(0.148)
σ_ϵ		4.063**
		(1.538)
σ_u		0.776**
		(0.131)
ρ_h		-0.349+
		(0.188)
Log-L	-5839.8544	-5558.0768
Observations	3414	
Individuals	957	

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Standard errors are estimated using BHHH (or OPG) estimator. Standard errors for reparameterized coefficients are estimated using Delta method.

Source: Author's estimation using 1985-2005 PSID data

Table 5: **Estimation Results for Individuals Who Have At Least 4 Years of Observations**

Variables	**No Individual Heterogeneity**	**Random Effects**	**Fixed Effects**
Child Support Payment Amount (1,000's)			
Constant	-5.141**	-4.346**	
	(0.398)	(0.702)	
Total Income (10,000's)	0.389**	0.351**	0.253**
	(0.050)	(0.064)	(0.082)
Father's Income (10,000's)	0.268**	0.121	0.031
	(0.082)	(0.111)	(0.117)
# of own children in the HH	-0.635**	-0.357+	-0.256
	(0.130)	(0.202)	(0.343)
# of other children in the HH	-0.408**	-0.388+	-0.079
	(0.126)	(0.222)	(0.359)
If father married again	1.428**	1.280**	0.839
	(0.289)	(0.464)	(0.544)
Years since previous marriage ended	-0.141**	-0.145**	-0.116*
	(0.029)	(0.043)	(0.058)
If two or more dependent children outside the HH	2.080**	1.459**	0.830+
	(0.232)	(0.369)	(0.477)
If father not white	-2.481**	-2.521**	
	(0.281)	(0.587)	
If father has college degree	1.195**	1.975**	
	(0.240)	(0.469)	
$y_2 = 1$ if does not comply with child support order			
Constant	1.042**	1.572**	
	(0.135)	(0.262)	
Total Income (10,000's)	-0.029+	-0.052*	-0.129**
	(0.016)	(0.026)	(0.038)
Father's Income (10,000's)	-0.042*	-0.013	0.070
	(0.021)	(0.038)	(0.053)
# of own children in the HH	0.116*	0.066	-0.040
	(0.047)	(0.086)	(0.136)
# of other children in the HH	0.102**	0.098	0.011
	(0.038)	(0.075)	(0.112)
If father married again	0.021	-0.016	-0.208
	(0.074)	(0.135)	(0.197)
Years since previous marriage ended	0.059**	0.072**	0.101**
	(0.009)	(0.017)	(0.029)

Continued on next page

Table 5 – continued from previous page

Variables	No Individual Heterogeneity	Random Effects	Fixed Effects
If two or more dependent	-0.121+	-0.100	0.096
children outside the HH	(0.062)	(0.108)	(0.173)
If father not white	0.353**	0.427+	
	(0.104)	(0.221)	
If father has college degree	0.060	0.033	
	(0.074)	(0.161)	
OCSE expenditures per	-0.119	-0.553+	-0.838*
single mother	(0.165)	(0.284)	(0.364)
If state has immediate	0.180	0.185	0.164
income withholding	(0.110)	(0.165)	(0.219)
If state adopted numerical	-0.203+	-0.217	-0.300
guidelines	(0.105)	(0.177)	(0.216)
If father self employed	0.366**	0.325+	0.246
	(0.111)	(0.179)	(0.226)
If father works in public sector	-0.325**	-0.499**	-1.077**
	(0.064)	(0.174)	(0.321)
σ_v	5.941**	4.488**	3.915**
	(1.864)	(0.768)	(0.870)
ρ	0.389	0.383+	0.308
	(0.268)	(0.196)	(0.243)
σ_ϵ		3.930*	
		(1.806)	
σ_u		0.828**	
		(0.164)	
ρ_h		-0.401*	
		(0.201)	
Log-L	-4122.0350	-3880.6680	-3177.8776
Observations	2385		
Individuals	377		

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Standard errors are estimated using BHHH (or OPG) estimator. Standard errors for reparameterized coefficients are estimated using Delta method.

Source: Author's estimation using 1985-2005 PSID data

Estimated coefficients in tables 4 and 5 measure the marginal effects of independent variables on the latent voluntary child support payment, y_1^*. Since fathers cannot make negative child support payments, marginal effects of independent variables on the actual voluntary payments, $\tilde{y_1} \equiv max\,(0, y_1^*)$, are of greater interest. Marginal effects of x on $\tilde{y_1}$, conditional on individual heterogeneity, ϵ, are given by:

$$\frac{\partial \mathbf{E}\left[\tilde{y_1}|\, x, \epsilon\right]}{\partial x} = \Phi\left(\frac{\hat{\beta}' x + \sigma_\epsilon \epsilon}{\sigma_\nu}\right)\hat{\beta}, \tag{24}$$

where individual subscripts are dropped for notational simplicity (see Cameron and Trivedi 2005, p. 542 for derivation). We can estimate unconditional marginal effects for each observation by taking expectation over ϵ of equation (24) and evaluating x's at their actual values[18]:

$$\frac{\partial \mathbf{E}_\epsilon\left[\mathbf{E}\left[\tilde{y_1}|\, x, \epsilon\right]\right]}{\partial x} \approx \sum_l W_l \Phi\left(\frac{\hat{\beta}' x + \sigma_\epsilon \epsilon_l}{\sigma_\nu}\right)\hat{\beta}, \tag{25}$$

where expectation over individual heterogeneity term is approximated using Gauss-Hermite quadrature with nodes ϵ_l and weights W_l[19]. Then average marginal effects are estimated by taking simple average over individual marginal effects. Marginal effects for specification with no heterogeneity are estimated by setting $\sigma_\epsilon = 0$ and for Fixed Effects regression by setting $\sigma_\epsilon = 1$ in equation (24).

Estimated marginal effects are reported in table 6. As the table indicates, \$10,000 dollar increase in father's household annual income, on average, raises voluntary child support payments by \$100 per year. If increase in household income was entirely because of higher father's income, average child support is higher by additional \$170, as suggested by the Pooled regression, or it is higher by \$120, according to the Random Effects Regression. Fixed effects specification suggests no differential effect of father's individual income. Child support

[18]One could also estimate marginal effects of observed child support payments, y, conditional on noncompliance (which would involve more complicated expressions); however, I am interested in voluntary child support payments behavior, and not in the observed payment amounts.

[19]Note, that we should actually use conditional $\hat{\beta\left(\epsilon_l\right)}$, i.e. we should estimate MLE $\hat{\beta}$ for each value of ϵ_l. I am planning to implement this correction in the next version of the paper.

Table 6: **Marginal Effects**

Variables	No Individual Heterogeneity	Random Effects	Fixed Effects
Child Support Payment Amount (1,000's)			
Total Income (10,000's)	0.09	0.09	0.10
Father's Income (10,000's)	0.17	0.12	0.01
# of own children in the HH	-0.18	-0.11	-0.10
# of other children in the HH	-0.15	-0.15	-0.03
If father married again	0.38	0.38	0.34
Years since previous marriage ended	-0.05	-0.05	-0.05
If two or more dependent children outside the HH	0.82	0.67	0.33
If father not white	-0.81	-0.82	
If father has college degree	0.55	0.79	
Person-years	3414	3414	2385
Individuals	957	957	377

Source: Author's estimation using 1985-2005 PSID data

payments decrease with additional children in the father's household and with additional years since divorce or separation. They are significantly higher if a father is white or has at least college degree. Finally, if father is married to his new partner, as opposed to just cohabiting, child support payments go up by almost $400 per year. This potentially indicates that fathers who are more responsible individuals and care about their children are more attractive marriage partners.

6 Concluding Remarks

I find that higher share of father's income in household income increases the child support payment amounts. This finding rejects income pooling and is consistent with Family Bargaining models. However, after controlling for unobserved individual heterogeneity in RE specification, the differential effect of father's income significantly declines, while FE specification suggests that distribution of individual incomes plays no role after controlling for total household income.

I hypothesize that the difference in RE and FE specification results suggests that permanent (or potential) and not transitory income influences spouses' bargaining power. I am planning to explore this hypothesis by including other indicators of permanent (potential) income in RE specification, such as spouses' relative age, education, or average income. On the other hand, if unobserved individual heterogeneity in father's preferences for his children's welfare is correlated with his productivity and thus his income, FE specification is more appropriate, since Pooled and RE estimates will be biased. In the latter case, I cannot reject income pooling, which suggests that the commonly used Unitary Household model might be an appropriate modeling choice.

However, this study has some important limitations. One of the weaknesses of the Maximum Likelihood estimation used in this paper is its heavy reliance on distributional assumption about the error terms, which results in inconsistent estimates if errors are heteroscedastic or nonnormal[20] (Cameron and Trivedi 2005, p. 538). Another, probably the most serious weakness of my analysis is the fact that I do not observe actual child support court orders and use guideline amount as a proxy for order amount. Thus, I might be misclassifying some cases when fathers pay voluntarily vs. just complying with court orders. Monte Carlo experiments suggest that such misclassification results in biased coefficients. I am planning to modify the log-likelihood function to allow for misclassification by modeling the measurement error in the "observed" court order and, if data allows identification, estimating its distribution. Finally, I could use actual guideline formulas (and not just predicted guideline amounts) to estimate child support court orders and thus get a better proxy of court order amounts.

[20] In the next version of the paper I could model heteroscedastic errors.

References

Aizer, A. and S. McLanahan (2006). The Impact of Child Support Enforcement on Fertility, Parental Investments, and Child Well-Being. *Journal of Human Resources XLI*(1), 28–45.

Apps, P. F. and R. Rees (1997). Collective labor supply and household production. *The Journal of Political Economy 105*(1), 178–190.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics - Methods and Applications*. Cambridge University Press, New York, NY.

Carro, J. M. (2007, October). Estimating dynamic panel data discrete choice models with fixed effects. *Journal of Econometrics 140*(2), 503–528.

Case, A. C., I.-F. Lin, and S. S. McLanahan (2003, February). Explaining trends in child support: Economic, demographic, and policy effects. *Demography 40*(1), 171–189.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica 56*(1), 63–90.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy 100*(3), 437–467.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002, February). Marriage market, divorce legislation, and household labor supply. *J POLIT ECON 110*(1), 37–.

Del Boca, D. and C. J. Flinn (1995, December). Rationalizing child-support decisions. *The American Economic Review 85*(5), 1241–1262.

Ermisch, J. and C. Pronzato (2008). Intra-household allocation of resources: Inferences from non-resident fathers' child support payments. *The Economic Journal 118*(527), 347–362.

Fortin, B. and G. Lacroix (1997). A test of the unitary and collective models of household labour supply. *The Economic Journal 107*(443), 933–955.

Freeman, R. B. and J. Waldfogel (2001). Dunning delinquent dads: The effects of child support enforcement policy on child support receipt by never married women. *The Journal of Human Resources 36*(2), 207–225.

Garfinkel, I., T. Heintze, and C.-C. Huang (2001). *The Incentives of Government Programs and The Well-Being of Families*, Chapter Child Support Enforcement: Incentives and Well-being. Chicago, IL: Joint Center for Poverty Research.

Grall, T. S. (2006). Custodial mothers and fathers and their child support: 2003. *Current Population Reports, U.S. Census Bureau* (P60-225).

Greene, W. H. (1998). *LIMDEP, version 7, User's Manual*. Econometric Software, Inc., Plainview.

Heckman, J. J. (1981). *Structural Analysis of Discrete Data with Econometric Applications*, Chapter The incidental parameters problem and the problem of initial conditions in estimating a discrete-time data stochastic process. MIT Press, New York.

Honore, B. E. (1993, September). Orthogonality conditions for tobit models with fixed effects and lagged dependent variables. *Journal of Econometrics 59*(1-2), 35–61.

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997). Do husbands and wives pool their resources? evidence from the united kingdom child benefit. *Journal of Human Resources 32*(3), 463–480.

Manser, M. and M. Brown (1980). Marriage and household decision-making: A bargaining analysis. *International Economic Review 21*(1), 31–44.

McElroy, M. B. and M. J. Horney (1981). Nash-bargained household decisions: Toward a generalization of the theory of demand. *International Economic Review 22*(2), 333–349.

Morgan, L. W. and M. C. Lino (1999, Spring). A comparison of child support awards calculated under states' child support guidelines with expenditures on children calculated by the u.s. department of agriculture. *Family Law Quarterly 33*(1), 191–218.

National Conference of State Legislatures (2007). Termination of child support and support beyond majority. *http://www.ncsl.org/programs/cyf/educate.htm, accessed 4/18/2007*.

National Women's Law Center (2002). Dollars and sense: Improving the determination of child support obligations for low-income mothers, fathers and children. *National Women's Law Center and the Center on Fathers, Families, and Public Policy*.

Page, M. E. and A. H. Stevens (2004). The Economic Consequences of Absent Parents. *J. Human Resources XXXIX*(1), 80–107.

Panel Study of Income Dynamics (2006). *1968-2005 Individual and Famly public use datasets*. Produced and distributed by the University of Michigan with primary funding from the National Science Foundation, the National Institute of Aging, and the National Institute of Child Health and Human Development. Ann Arbor, MI.

Peterson, J. L. and C. W. Nord (1990). The regular receipt of child support: A multistep process. *Journal of Marriage and the Family 52*(2), 539–551.

Pirog, M. A., M. E. Klotz, and K. V. Byers (1998, jul). Interstate comparisons of child support orders using state guidelines. *Family Relations 47*(3), 289–295.

Smock, P. J. and W. D. Manning (1997). Nonresident parents' characteristics and child support. *Journal of Marriage and the Family 59*(4), 798–808.

Sorensen, Elaine and Halpern, Ariel (1999). Child support enforcement: How well is it doing? *Urban Institute Discussion Paper 99*(11).

Venohr, J. C. and T. E. Griffith (2005, July). Child support guidelines: Issues and reviews. *Family Court Review 43*(3), 415–428.

Weiss, Y. and R. Willis (1989). An economic analysis of divorce settlements. *University of Chicago - Economics Research Center Working Paper* (89-5).

Weiss, Y. and R. J. Willis (1985). Children as collective goods and divorce settlements. *Journal of Labor Economics 3*(3), 268–292.

Appendix A. Solving Father's Household Utility Maximization Problem

The father and his new partner maximize family's welfare function, which is a weighted sum of individual spouses' utilities, subject to a pooled income budget constraint and mother's expenditures on child quality:

$$\begin{aligned} \max_{c_f,c_p,t} U_f + \mu U_p &= \delta_f \log(c_f) + (1-\delta_f)\log(k) - \vartheta I[t<s] + \mu \log(c_p), \\ s.t. \quad & y_f + y_p = t + c_f + c_p, \\ & k = (1-\delta_m)(y_m + t). \end{aligned} \tag{26}$$

Denote the sum of mother's and father's household income as $y_T = y_m + y_f + y_p$. Then, assuming internal solution (i.e. assuming that noncompliance cost is low enough and father's preference towards child quality is high enough), we get the following optimal consumption and child support transfer amounts:

$$\begin{aligned} c_f^* &= \tfrac{\delta_f}{1+\mu} y_T, \\ c_p^* &= \tfrac{\mu}{1+\mu} y_T, \\ t^* &= \tfrac{1-\delta_f}{1+\mu}(y_f + y_p) - \tfrac{\mu+\delta_f}{1+\mu} y_m. \end{aligned} \tag{27}$$

If father decides to avoid noncompliance cost and complies with child support order by paying $t = s$, then mother's expenditures on child quality are given by $k^{EC} = (1-\delta_m)(y_m + s)$, while father's and his partner's optimal consumption amounts are found by solving the following maximization problem:

$$\begin{aligned} \max_{c_f,c_p} U_f + \mu U_p &= \delta_f \log(c_f) + (1-\delta_f)\log((1-\delta_m)(y_m+s)) + \mu \log(c_p), \\ s.t. \quad & y_f + y_p = t + c_f + c_p, \end{aligned} \tag{28}$$

I refer to this situation as "Exact compliance" case and it results in the following consumption levels:

$$\begin{aligned} c_f^{EC} &= \tfrac{\delta_f}{1+\mu} \left(y_f + y_p - s \right), \\ c_p^{EC} &= \tfrac{\mu}{1+\mu} \left(y_f + y_p - s \right). \end{aligned} \tag{29}$$

In the case of "Over Compliance", father pays more child support than the court order amount. This happens when $t^* (\delta_f) > s$, or $\delta_f < 1 - \frac{(1+\mu)(y_m+s)}{y_T} \equiv \underline{\delta}$. In this case, the level of noncompliance cost ϑ is irrelevant to father's decision problem. When the father pays positive child support which is lower than the court order, we have a "Partial Payments" case. Finally, the father voluntarily pays no child support if $t^* (\delta_f) \leq 0$, or $\delta_f > 1 - (1 + \mu) \frac{y_m}{y_T} \equiv \overline{\delta}$. I call this situation as the "No Payments" case. In the latter case father's and his partner's optimal consumption levels are given by:

$$\begin{aligned} c_f^{NP} &= \tfrac{\delta_f}{1+\mu} \left(y_f + y_p \right), \\ c_p^{NP} &= \tfrac{\mu}{1+\mu} \left(y_f + y_p \right). \end{aligned} \tag{30}$$

When father's voluntary child support transfer amount is less than the court order, the father has to decide whether he should not comply with the court order and incur noncompliance cost, or whether he should comply with the court order and have suboptimal consumption levels. He can make this decision by comparing his household's utility function values in both cases. Denote his and his partner's household's indirect utility level in the case of "Exact compliance" as $W^{EC} = \delta_f \log \left(c_f^{EC} \right) + (1 - \delta_f) \log \left(k^{EC} \right) + \mu \log \left(c_p^{EC} \right)$. Moreover, denote father's household's indirect utility excluding noncompliance cost term in the case of "No Payments" as $W^{NP} = \delta_f \log \left(c_f^{NP} \right) + (1 - \delta_f) \log \left((1 - \delta_m) \, y_m \right) + \mu \log \left(c_p^{NP} \right)$, and in the case of "Partial Payments" as $W^{PP} = \delta_f \log \left(c_f^* \right) + (1 - \delta_f) \log \left((1 - \delta_m) \left(y_m + t^* \right) \right) + \mu \log \left(c_p^* \right)$. Then father decides to comply with the court order if noncompliance cost is high enough, i.e. if $W^{EC} > W^{NP} - \vartheta$ given that $\delta_f > \overline{\delta}$, or if $W^{EC} > W^{PP} - \vartheta$ given that $\underline{\delta} < \delta_f \leq \overline{\delta}$.

Therefore, the solution of household's utility maximization problem can be separated into four cases, as defined above, depending on the values of father's preference and noncompliance cost parameters:

$$\begin{array}{llll} 1)\ No\ Payments & t=0 & if\ \delta_f \in \left(\overline{\delta},1\right]\ and\ \vartheta \in \left[0, W^{NP}-W^{EC}\right), \\ 2)\ Partial\ Payments & t=t^{*}<s & if\ \delta_f \in \left(\underline{\delta},\overline{\delta}\right]\ and\ \vartheta \in \left[0, W^{PP}-W^{EC}\right), \\ 3)\ Exact\ Compliance & t=s & if \begin{array}{l} \delta_f \in \left(\overline{\delta},1\right]\ and\ \vartheta \in \left[W^{NP}-W^{EC},\infty\right) \\ \delta_f \in \left(\underline{\delta},\overline{\delta}\right]\ and\ \vartheta \in \left[W^{PP}-W^{EC},\infty\right) \end{array} \\ 3)\ Over\ Compliance & t=t^{*}>s & if\ \delta_f \in \left[0,\underline{\delta}\right], \end{array} \tag{31}$$

Appendix B. Individual Likelihood and Gradient Specifications

Density function for each observation depends on actual child support payment and court order amounts and can be decomposed into four different parts as described in equation (7). Density function for these four parts, conditional on individual heterogeneity terms, ϵ_i and u_i, is the following:

1. When observed $y_{1it}=0$:

$$\begin{aligned} f_1\left(y_{it}|\,u_i,\epsilon_i\right) &= Pr\left(y_{1it}^{*}\leq 0, y_{2it}^{*}>0\right) = Pr\left(\tfrac{\nu_{it}}{\sigma_\nu}\leq \tfrac{-\beta' x_{it}-\sigma_\epsilon \epsilon_i}{\sigma_\nu}, \omega_{it} > -\gamma' z_{it} - \sigma_u u_i\right) \\ &= Pr\left(\tfrac{\nu_{it}}{\sigma_\nu}\leq \tfrac{-\beta' x_{it}-\sigma_\epsilon \epsilon_i}{\sigma_\nu}, -\omega_{it} \leq \gamma' z_{it} + \sigma_u u_i\right) \\ &= \Phi_2\left(\tfrac{-\beta' x_{it}-\sigma_\epsilon \epsilon_i}{\sigma_\nu}, \gamma' z_{it}+\sigma_u u_i, -\rho\right), \end{aligned} \tag{32}$$

since Gaussian distribution is symmetrical. Here Φ_2 denotes bivariate standard normal CDF.

2. When $0 < y_{1it} < s_i$:

$$\begin{aligned} f_2\left(y_{it}|\, u_i, \epsilon_i\right) &= Pr\left(0 < y_{1it}^* < s_i, y_{2it}^* > 0\right) f\left(y_{1it}^* | 0 < y_{1it}^* < s_i, y_{2it}^* > 0\right) \\ &= Pr\left(0 < y_{1it}^* < s_i, y_{2it}^* > 0 | y_{1it}^*\right) f\left(y_{1it}^*\right) \\ &= Pr\left(y_{2it}^* > 0 | y_{1it}^*\right) f\left(y_{1it}^*\right) \\ &= \Phi\left(\frac{\gamma' z_{it} + \sigma_u u_i + \frac{\rho}{\sigma_\nu}\left(y_{it} - \beta' x_{it} - \sigma_\epsilon \epsilon_i\right)}{\left(1-\rho^2\right)^{1/2}}\right) \frac{1}{\sigma_\nu} \phi\left(\frac{y_{it} - \beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu}\right), \end{aligned} \tag{33}$$

since $\omega_{it}|_{\nu_{it}} = \frac{\rho}{\sigma_\nu}\nu_{it} + \xi_{it}$, $\xi_{it} = N\left(0, 1-\rho^2\right)$, where Φ stands for standard normal CDF and ϕ denotes standard normal PDF.

3. When $y_{1it} = s_i$:

$$\begin{aligned} f_3\left(y_{it}|\, u_i, \epsilon_i\right) &= Pr\left(y_{1it}^* \leq s_i, y_{2it}^* \leq 0\right) = Pr\left(\frac{\nu_{it}}{\sigma_\nu} \leq \frac{s_i - \beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu}, \omega_{it} \leq -\gamma' z_{it} - \sigma_u u_i\right) \\ &= \Phi_2\left(\frac{s_i - \beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu}, -\gamma' z_{it} - \sigma_u u_i, \rho\right) \end{aligned} \tag{34}$$

4. Finally, when $y_{1it} > s_i$:

$$\begin{aligned} f_4\left(y_{it}|\, u_i, \epsilon_i\right) &= Pr\left(y_{1it}^* > s_i\right) f\left(y_{1it}^* | y_{1it}^* > s_i\right) = f\left(y_{1it}^*\right) \\ &= \frac{1}{\sigma_\nu} \phi\left(\frac{y_{it} - \beta' x_{it} - \sigma_\epsilon \epsilon_i}{\sigma_\nu}\right) \end{aligned} \tag{35}$$

In order to simplify notation, define the following:

$$\begin{gathered} \tau_1 \equiv \sigma_\epsilon = \exp\left(\tfrac{1}{2}\alpha_\epsilon\right);\ \tau_2 \equiv \sigma_u = \exp\left(\tfrac{1}{2}\alpha_u\right); \\ \delta_1 \equiv \rho = \frac{1-\exp(\alpha_\rho)}{1+\exp(\alpha_\rho)};\ \delta_2 \equiv \frac{1}{\left(1-\rho^2\right)^{1/2}} = \frac{1+\exp(\alpha_\rho)}{2\exp\left(\frac{1}{2}\alpha_\rho\right)};\ \delta_3 \equiv \frac{1}{\sigma_\nu} = \exp\left(-\tfrac{1}{2}\alpha_\nu\right); \\ A_{1it} \equiv \frac{\beta' x_{it} + \sigma_\epsilon \epsilon_i}{\sigma_\nu} = \delta_3\left(\beta' x_{it} + \tau_1 \epsilon_i\right);\ A_{2it} \equiv \delta_3\left(y_{it} - \beta' x_{it} - \tau_1 \epsilon_i\right); \\ A_{3it} \equiv \delta_3\left(\beta' x_{it} + \tau_1 \epsilon_i - s_i\right);\ B_{it} \equiv \gamma' z_{it} + \tau_2 u_i. \end{gathered}$$

Then the expression for the log-likelihood for each observation is:

$$\begin{aligned} l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right) &\equiv \log\left(f\left(y_{it}|\,u_m,\epsilon_l\right)\right) = I\left(y_{it}=0\right)\times\left[\log\Phi_2\left(-A_{1it},B_{it},-\delta_1\right)\right] \\ &+I\left(0<y_{it}<s_i\right)\times\left[\log\Phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)\right. \\ &\left.+\log\left(\delta_3\right)-\tfrac{1}{2}\log\left(2\pi\right)-\tfrac{1}{2}A_{2it}^2\right] \\ &+I\left(y_{it}=s_i\right)\times\left[\log\Phi_2\left(-A_{3it},-B_{it},\delta_1\right)\right] \\ &+I\left(y_{it}>s_i\right)\times\left[\log\left(\delta_3\right)-\tfrac{1}{2}\log\left(2\pi\right)-\tfrac{1}{2}A_{2it}^2\right] \end{aligned} \tag{36}$$

The remaining of this section specifies the expressions for $\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)/\partial\tilde{\theta}$ for each of the four cases defined in equation (7).

1. For the case of no child support payment, i.e. when $y_{1it}=0$:

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial\left[\begin{array}{ccc}\beta' & \alpha_\epsilon' & \alpha_\nu'\end{array}\right]'} = -\delta_3\frac{\partial l_{1it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\Phi_2\left(-A_{1it},B_{it},-\delta_1\right)}\left[\begin{array}{c} x_{it} \\ \frac{1}{2}\tau_1\epsilon_i \\ -\frac{1}{2}A_{1it}/\delta_3\end{array}\right]$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial\left[\begin{array}{cc}\gamma' & \alpha_u'\end{array}\right]'} = \frac{\phi\left(B_{it}\right)\Phi\left(\delta_2\left(-A_{1it}+\delta_1 B_{it}\right)\right)}{\Phi_2\left(-A_{1it},B_{it},-\delta_1\right)}\left[\begin{array}{c} z_{it} \\ \frac{1}{2}\tau_2 u_i\end{array}\right]$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial\alpha_\rho} = \frac{1}{2}\delta_2^{-2}\frac{\phi_2\left(-A_{1it},B_{it},-\delta_1\right)}{\Phi_2\left(-A_{1it},B_{it},-\delta_1\right)}$$

2. When $0<y_{1it}<s_i$:

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial\left[\begin{array}{cc}\beta' & \alpha_\epsilon'\end{array}\right]'} = -\delta_3\left(\delta_1\delta_2\frac{\phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}{\Phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}-A_{2it}\right)\left[\begin{array}{c} x_{it} \\ \frac{1}{2}\tau_1\epsilon_i\end{array}\right]$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial\left[\begin{array}{cc}\gamma' & \alpha_u'\end{array}\right]'} = \delta_2\frac{\phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}{\Phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}\left[\begin{array}{c} z_{it} \\ \frac{1}{2}\tau_2 u_i\end{array}\right]$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{=} - \frac{1}{2}\delta_1\delta_2\frac{\phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}{\Phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}A_{2it}-\frac{1}{2}+\frac{1}{2}A_{2it}^2$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial\alpha_\rho} = -\frac{1}{2}\delta_2\frac{\phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}{\Phi\left(\delta_2\left(B_{it}+\delta_1 A_{2it}\right)\right)}\left(\left(\delta_1^2+\delta_2^{-2}\right)A_{2it}+\delta_1 B_{it}\right)$$

3. When $y_{1it} = s_i$:

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial \left[\begin{array}{ccc}\beta' & \alpha'_\epsilon & \alpha'_\nu\end{array}\right]'} = -\delta_3 \frac{\phi\left(A_{3it}\right)\Phi\left(\delta_2\left(-B_{it}+\delta_1 A_{3it}\right)\right)}{\Phi_2\left(-A_{3it},-B_{it},\delta_1\right)} \left[\begin{array}{c} x_{it} \\ \frac{1}{2}\tau_1\epsilon_i \\ -\frac{1}{2}A_{3it}/\delta_3 \end{array}\right]$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial \left[\begin{array}{cc}\gamma' & \alpha'_u\end{array}\right]'} = -\frac{\phi\left(B_{it}\right)\Phi\left(\delta_2\left(-A_{3it}+\delta_1 B_{it}\right)\right)}{\Phi_2\left(-A_{3it},-B_{it},\delta_1\right)} \left[\begin{array}{c} z_{it} \\ \frac{1}{2}\tau_2 u_i \end{array}\right]$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial \alpha_\rho} = -\frac{1}{2}\delta_2^{-2}\frac{\phi_2\left(-A_{3it},-B_{it},\delta_1\right)}{\Phi_2\left(-A_{3it},-B_{it},\delta_1\right)}$$

4. Finally, when $y_{1it} > s_i$:

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial \left[\begin{array}{cc}\beta' & \alpha'_\epsilon\end{array}\right]'} = \delta_3 A_{2it} \left[\begin{array}{c} x_{it} \\ \frac{1}{2}\tau_1\epsilon_i \end{array}\right]$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial \left[\begin{array}{ccc}\gamma' & \alpha'_u & \alpha'_\rho\end{array}\right]'} = \mathbf{0}$$

$$\frac{\partial l_{it}\left(\tilde{\theta}\,|u_m,\epsilon_l\right)}{\partial \alpha_\nu} = -\frac{1}{2} + \frac{1}{2}A_{2it}^2$$

Appendix C. Hessian of the concentrated log-likelihood in Fixed Effects Estimation

Expression for Hessian of the concentrated log-likelihood is given by

$$\begin{aligned}\frac{\partial^2 l^C(\theta)}{\partial\theta\partial\theta'} &= \sum_i \Big[d_{\theta\theta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) + 2 d_{\theta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) \frac{\partial\hat{\eta}_i(\theta)}{\partial\theta'} \\ &+ \frac{\partial\hat{\eta}_i(\theta)'}{\partial\theta} d_{\eta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) \frac{\partial\hat{\eta}_i(\theta)}{\partial\theta'} + d_{\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) \frac{\partial^2\hat{\eta}_i(\theta)}{\partial\theta\partial\theta'} \Big]\end{aligned} \tag{37}$$

This can be simplified by noting, that $d_{\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) \equiv 0$, which we can differentiate w.r.t. θ:

$$d_{\theta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right)' + d_{\eta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) \frac{\partial\hat{\eta}_i\left(\theta\right)}{\partial\theta'} \equiv 0, \tag{38}$$

or

$$\frac{\partial\hat{\eta}_i\left(\theta\right)}{\partial\theta'} \equiv -\left[d_{\eta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right)\right]^{-1} d_{\theta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right)' \tag{39}$$

Then after substituting for $\frac{\partial\hat{\eta}_i(\theta)}{\partial\theta'}$ and using the fact that $d_{\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) \equiv 0$, the final expression for Hessian becomes the following:

$$\frac{\partial^2 l^C\left(\theta\right)}{\partial\theta\partial\theta'} = \sum_i \left[d_{\theta\theta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) - d_{\theta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right) \left[d_{\eta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right)\right]^{-1} d_{\theta\eta i}\left(\theta, \hat{\eta}_i\left(\theta\right)\right)' \right], \tag{40}$$

where $d_{\theta\theta i}\left(\theta, \eta_i\right) \equiv \sum_t \frac{\partial^2 l_{it}(\theta,\eta_i)}{\partial\theta\partial\theta'}$, $d_{\theta\eta i}\left(\theta, \eta_i\right) \equiv \sum_t \frac{\partial^2 l_{it}(\theta,\eta_i)}{\partial\theta\partial\eta_i'}$ and $d_{\eta\eta i}\left(\theta, \eta_i\right) \equiv \sum_t \frac{\partial^2 l_{it}(\theta,\eta_i)}{\partial\eta_i\partial\eta_i'}$ are estimated by numerically differentiating score functions.