The income earned by women: impacts on welfare outcomes

Lawrence Haddad

International Food Policy Research Institute 2033 K Street, NW, Washington DC 20006, USA

Received 19 November 1997; received in revised form 21 July 1998; accepted 29 October 1998

Abstract

This paper provides an updated review of the evidence on income pooling across household members. Income pooling is one of the main predictions of the unitary model of the household. New studies come to much the same conclusion as do past studies: income pooling and the unitary model are rejected. The paper then looks beyond the mere rejection of the unitary model and explores some of the issues that arise. First, what is the progress in testing the restrictions imposed by non-unitary models of the household? Second, what are the implications of rejection of the unitary model for policy and program design? Finally, what are some of the challenges faced by programs and policies that internalize the rejection of income pooling in terms of impact evaluation? © 1999 Elsevier Science B.V. All rights reserved.

1. Introduction

There is an increasingly convincing set of studies in the economics and sociology literature which suggests that the marginal effect of income in the hands of women is different from income in the hands of men. This result implies incomplete income pooling within the household and is a refutation of a model of intra-household resource allocation that would have us believe that household members maximize a single welfare function, the so-called 'unitary' model of the household (Alderman et al., 1995; Behrman, 1997; Bergstrom, 1997; Haddad et al., 1997).

This paper seeks to add to this literature in a number of modest ways. First, we briefly review and update the evidence on income source dependence. What have recent studies had to say about income pooling and the validity of the unitary model? The vast amount of attention given to the refutation of the unitary model has not been matched by a discussion of the issues that emerge above and beyond rejection of the unitary model of the household. In this paper we touch upon three areas that need more work: (1) efforts to test models that are alternatives to the unitary model, (2) the implications for programs and policy design of the non-unitary household, and (3) the potential for randomized studies that address income source dependence in the context of program evaluation.

2. Evidence on the differential marginal impacts of male and female income

The idea that men and women spend income from own-earnings in different ways is not new. Together, papers by Kumar (1979); Guyer (1980); Tripp (1981), and Pahl (1983) document this phenomenon over a wide range of settings and times. Why do men and women tend to spend income differently? Societal and
cultural norms may assign women the role of ‘gatekeepers,’ in which they ensure that household members, especially children, receive an adequate share of available food. Alternatively, women may prefer to spend more on children’s daily needs because they spend more time with them. Women may also face different constraints than men. To minimize the competing demands on their time, for example, women may spend more on food because they purchase more expensive calories that take less time to prepare. Finally, women and men may have different income flows and thus different transaction costs. In other words, since women’s income tends to come more frequently and in smaller amounts, it may be more readily spent on household daily subsistence needs than lumpier seasonal income, which tends to come to men and is likely to be spent on more expensive items (Hamilton et al., 1984).

It is important to note that the observation that men and women spend income from own-earnings in different ways is fully consistent with both the standard unitary (or common preference or household utility or benevolent dictator or altruistic) model of the household and collective (or non-unitary) models of household resource allocation. As stated above, it may be the case, for example, that men and women simply earn income in flows and forms in which it makes sense for women to purchase food and men to purchase non-foods. In this case the decision to purchase food and the decision to undertake a particular type of employment are simultaneously determined. This is an example of endogeneity of male and female income with the outcome of interest – food expenditures in this instance.

The observation that men and women spend marginal income from own-earnings in different ways may also be consistent with the unitary model due to statistical measurement error. If, for example, female income streams are measured with more random error than male income streams, marginal propensity estimates that fail to adequately compensate for such errors will be different for male and female income. In order to refute the unitary model via income pooling and similar tests it is necessary to address possible endogeneity and measurement error.

2.1. Attempts to deal with endogeneity and measurement error

To reject the hypothesis that the marginal propensity to consume out of the income earned by men and women is identical is only a rejection of the unitary model when issues of endogeneity and measurement error are convincingly dealt with.

Studies by Schultz (1990) and Thomas (1990, 1992) were among the first to deal with the issue of endogeneity of income. As (Schultz, 1990, pp. 601–602) notes “If non-earned income (or ownership of the underlying asset) influences family demand behavior differently, depending on who in the family controls the income (or owns the asset), then the preferences for that demand must differ across individuals and such families must not completely pool unearned income.” Using non-labor income, Schultz (1990), and Thomas (1990, 1992) find that increased (non-labor) income received by women leads to a greater share of the household budget devoted to expenditures on human capital and a higher level of nutrient intake.

The use of non-labor income does ameliorate the problem of endogeneity of income, but it has nevertheless been subject to criticism because it is likely to reflect past labor allocation choices (Haddad et al., 1997). The approach has also come under scrutiny because unearned income is typically a small component of income and one that may be less likely to be measured accurately. By examining the impacts of male and female income on boy and girl welfare outcomes Thomas (1994, 1997) claims that errors in variables cannot explain the different impacts of female income on daughters and sons.

Focusing on labor income probably reduces problems of measurement error, but the credibility of the estimates depends on finding credible instruments for male and female income. Hoddinott and Haddad (1995) use traditional cropping patterns in Cote D’Ivoire to provide identification of income sources and by instrumenting the share of household income

---

2One study has attempted to test this hypothesis by controlling for the flows of incomes earned by men and women in Niger (Hopkins et al., 1994). Conditional on the study being only partially successful in controlling for endogeneity of separate incomes with instrumental variables, the findings indicate that the level and timing of female income flows has a significant effect on both total household expenditures and food expenditures in a given season, while the timing of male income flows has no effect. Thus both flow and gender of income earner matter.
earned by women they find that a higher share of income to women leads to higher expenditures on food and lower expenditures on alcohol and cigarettes. Haddad and Hoddinott (1994) use the same data to demonstrate differential impacts of women's income on child anthropometry outcomes.

Rogers et al. (1996) conduct a similar study using data from Honduras. They find that “controlling for income level, the percent of income earned by the child’s mother has a positive effect on nutritional outcome ... the positive effect ... decreases as children get older” (p.150). A crucial problem with this study, however, is that it does not instrument for the percent of income earned by mothers. Using data from Zambia, Wang (1996) uses fixed-effects instrumental variables methods to control for the endogeneity of mothers’ income. The paper finds that mothers’ income increases infant girls’ weight for age more than it does for infant boys (difference significant at the 10% level) and that fathers’ income increases boys’ weight and height for age more than for girls (significant at the 5% level). The problem here is that the use of child fixed effects means we can only interact female and male income with child outcomes, thus giving us only the differential effects of these incomes on boy and girl outcomes (see Haddad and Hoddinott, 1994 for a similar approach).

Doss (1997), makes a virtue out of a lack of individual income streams in her Ghana data by using assets owned by different individuals in the household. Using ordinary least squares she estimates a series of budget share equations. Controlling for overall asset levels, she finds that, for urban households, female asset share has a significant impact on seven out eight budget share categories: food (+), alcohol (-), education (+), recreation (-), tobacco (-), household (+), misc (-); female farmland share has a significant impact on food (+), alcohol (-), recreation (-), tobacco (-), and completeness of vaccinations (+). For rural households the female asset share significantly affects four out of eight budget share categories: food (+), alcohol (-), recreation (-) and tobacco (-), and child education outcomes (+). Doss concludes that it is important not to treat the household as if it made decisions based on a unitary model. Nevertheless, one could claim that even current asset holdings are not exogenous, and that these should be instrumented by information such as parental assets or assets brought to marriage (Quisumbing, 1994). Thomas et al. (1997) do in fact employ this approach to instrument for non-labor (or asset) income.

Another approach to avoiding endogeneity of income source and level is to look for regime switches that affect exogenous sources of income. Work by Lundberg et al. (1995) and Ward-Batts (1996) use data from the United Kingdom for two periods (1973–1976 and 1980–1982) attempts to assess the impact on the budget shares of 23 goods of a policy shift in U.K. child benefit allocation. The legislative change resulted in income transfers being directed specifically to women in the form of cash as opposed to a deduction from the household’s income tax. Using a dummy variable for policy regime, the study done by Ward-Batts found that the legislative change had a strong negative impact on budget shares to tobacco, housing, and men’s clothing and a strong positive impact on children’s clothing, fuel, and food purchased for home consumption. Lundberg et al. (1997) find similar results in that expenditures on women’s and child clothing increased relative to men’s clothing expenditures after the policy change.

Hoddinott and Adam (1998) capitalize on a change in Canadian state law regarding the dispensation of income and assets upon divorce. The law change improved the likelihood of women receiving a larger share of such resources. Using female suicide rates as the welfare outcome, the study found a significant drop in female suicide rates immediately after the regime change, which given the way the test was set up, as a Nash-bargaining model versus a unitary model, is a refutation of the latter.

Rubalcava and Thomas (1997) use inter-state and inter-temporal variation in Aid to Families with Dependent Children (AFDC) to test the unitary model of the household. They test whether the variation in the generosity of AFDC transfers, controlling for overall household income affects the share of expenditures on food. As AFDC benefits increase, the assumption is that this improves the fallback position of women and makes her better able to separate from the household if she is unable to negotiate what she considers a fair share of resources. Rubalcava and Thomas (1997) find that AFDC benefits do affect food expenditure shares and they conclude that their results ‘sink one more nail into the coffin of the unitary model of the household’ (p.20).
3. Beyond rejecting the unitary model

It is not exactly news that the unitary model has credibility problems.\(^3\) It is reassuring to find the latest evidence confirming the earlier studies. But what next? What are the implications for the next wave of model testing? For policy design? For policy evaluation? This section discusses these issues.

3.1. Testing alternatives

Chiappori (1997) notes that while evidence against income pooling may well signal a problem with the unitary model, this mere fact does not support any alternative model of household behavior in particular (how legitimate is it, for example, for Rubalcava and Thomas to motivate their empirical study with a Nash-bargained model; a model that is only one alternative to the unitary model?).\(^4\) The only way to empirically support a particular collective setting is to derive from that collective framework itself, conditions that can potentially be, but are actually not falsified by empirical observation. Imposing only Pareto efficiency on his collective model of household resource allocation Chiappori derives restrictions that if rejected will result in the model being rejected. The restrictions derived are that the ratio of the impact of A’s income upon the demand for Commodity i and the impact of B’s income upon the demand for Commodity i should be identical for all goods. These restrictions can be tested by incorporating them into a demand function that permits the coefficients on income from individual A and individual B to vary by commodity. Using a sample of French households in which both the husband and wife work full time and nine consumption goods, the restrictions implied by the collective model could not be rejected by the data at the 5% level.

Work by Bourguignon et al. (1993) and Browning et al. (1994) builds on Chiappori’s work, but still using the assumption of exogenous labor supply, an assumption that is particularly unrealistic in a developing country setting.

---

\(^3\) Although the unitary model might do violence to reality we are well-served to remember that it is still useful for examining intrahousehold issues (Haddad et al., 1997).

\(^4\) See McElroy (1997) for a discussion of different types of collective models.

3.2. Operationalizing the results

What are the policy and programmatic implications of differential impacts of women’s income? The results have implications in a number of settings. The implications in terms of the design of transfer programs are straightforward: when the benefits to targeting women (in terms of the intervention’s objective) outweigh the costs, then this should be undertaken.

It should be said that a number of programs and interventions are well ahead of the academics in terms of targeting income transfers to women. The U.K. child benefit example cited earlier is one such example. The practice of the large Bangladeshi NGOs such as the Grameen Bank, PROSHIKA and ASA to target credit to women is another. Pitt and Khandker (1996) use data from Bangladesh to test for a differential impact of male and female borrowing from NGOs on eight outcomes: boy’s and girl’s schooling, women’s and men’s labor supply, total household expenditure, contraception use, fertility, and the value of women’s non-land assets. Using a sophisticated econometric approach (Weighted Exogenous Sampling Maximum Likelihood-Fixed Effects) they take care to be able to identify the average impact of the credit program on the outcomes of interest. The key to their approach is to sample households in NGO program villages that are ineligible due to some exogenous constraint (such as an inability to meet the land-holding criteria due to the absence of a land market). They find that “the set of female credit variables is statistically significant in 7 of 8 cases at the 0.05 percent level. By contrast the set of male credit variables is significant in three out of eight cases... the hypothesis that male and female credit parameters are jointly equal for each of these
three (NGO) programs is rejected in only four cases: women’s labor supply, women’s non-land assets, contraception and fertility” (p.41). Taken at face value their results imply that loans to women have greater marginal impacts than loans to men. This is consistent with the refutation of the unitary model of the household, but it is also consistent with lack of full control for unobserved heterogeneity and with declining marginal returns to capital, since loans to men are larger than loans to women (Morduch, 1998).

In Mexico a large new program called PROGRESA began operation in August 1997 to fight ‘extreme poverty’ in Mexico’s rural areas. This multisectoral program provides an integrated package of health, nutrition, and educational services to poor families. Currently serving 400,000 families (Gomez de Leon et al., 1997), it aims to expand its coverage to approximately 1–1.5 million families by the end of 1998, with an approximate budget of 500 million dollars. PROGRESA is one of the Mexican government’s primary weapons against poverty. The program aims to provide a series of interventions, including monetary assistance, nutritional supplements, educational grants, and a basic health package, to its beneficiaries for at least 3 consecutive years. One of the innovative aspects of the program is its attempt to transfer the monetary assistance to women. The literature on the differential impacts of male and female income was influential in this aspect of the programs’ design.

The challenge is now to evaluate the benefits from targeting women. This is important because the attempt to target women is not costless. A number of potential unanticipated and costly consequences could arise from programs targeted in this way. In the context of the Grameen Bank Goetz and Sen Gupta (1994) find evidence of women acting as a ‘front’ for men who want to gain access to credit. This might increase transactions costs above those facing beneficiaries in an untargeted program, and may place women at risk of abuse.

3.3. The need for randomization studies

Despite the expanding literature on gender differences in the control of income, the evidence is all based on regression analyses of observational data. No matter how sophisticated the econometric techniques, the lingering doubts of the staunchest skeptics (in the academic and operational fields) will only be vanquished by an experimental design that involves a randomization of households into two groups, one group where income transfers are targeted to women and one group where they are targeted to men. In part to answer the skeptics and in part to evaluate new interventions that do attempt to target women, a new wave of studies might emerge that use the random assignment of income transfers to men and to women.

Randomization is attractive in theory, although the pathways behind the particular result, and the magnitude of the difference in household welfare outcomes between the two groups, are hard to uncover (Heckman and Smith, 1995). In practice, however, randomization may not prove to be so straightforward. In a true experimental evaluation where the treatment and control groups have been randomized and the evaluation is conducted in a double blind manner (investigators and beneficiaries are not aware of who is in the treatment and control groups), isolating the true impact of an intervention is fairly straightforward. It is of course not possible to conduct a double blind evaluation of an income transfer program (e.g., consider the problem of constructing a placebo for the control group).

There are additional problems that need to be addressed: (1) agreement and compliance with implementors, governments or NGOs, may prove to be difficult (e.g. implementors may be under extreme pressure to get the intervention to as many recipients as possible and may not want to delay the delivery of benefits to a randomized group for the purposes of evaluation); (2) ethical issues also need to be discussed: if our priority is that children will be better off with transfers targeted to women, are we comfortable with the experiment? (3) will the conduct of the experiment change behavior? (e.g. will men in the experiment consume less cigarettes and alcohol than they would if they were not participating in the experiment?); and (4) what to do in households in which adults are either all males or all females? Answers to these questions may well come from the biological sciences where such randomized intervention trials are standard tools for drawing inferences (see Gehlbach, 1982 and Esrey et al., 1985 for good introductions to these issues).

The PROGRESA program described above provides a good example of the difficulties of randomiza-
tion in the context of an operating program. In order to evaluate the impact of gender-targeting, it is necessary for different regions and communities to receive different packages. In a pilot study this may not be too much of a problem (see Garcia and Pinstrup-Andersen, 1987 for an example from the Philippines), but for a rapidly expanding program, that is even moderately politicized this may well be impossible. Given the circumstances under which PROGRESA is currently operating the manufacture of variation through randomizing the targeting of benefits delivered to different groups, perhaps on a rotating basis, seems to be too complicated administratively and too difficult politically. Are there any other approaches that can be used in situations like this? One option is to resort to methods that rely on the answers to a series of questions about hypothetical situations. The contingent valuation literature provides a number of tools that may be candidates for such an approach (Bishop et al., 1995).

4. Conclusions

The studies cited in this paper vary widely in terms of econometric rigor and sampling sophistication. Nevertheless, they are consistent in that they find some evidence of women’s share of income (or assets or credit) influencing a range of household outcomes, controlling for overall household resources. Despite the wide range of estimation techniques used, the studies demonstrate a pattern of rejection of the hypothesis that male and female income (or credit or assets) have equal marginal effects on a range of household and individual welfare outcomes. The more immediate implications of this result have already begun to be acted upon as the experiences from Bangladesh and Mexico illustrate. It will be important to measure the immediate impacts of such designs on household expenditures and child nutrition, and we have discussed some of the problems of such evaluations.

More importantly perhaps will be the medium-run impacts of such interventions on the status of women within the household and, in turn, the impact of improved status on a number of important policy objectives: fertility reduction, food purchases, nutrition, and education to name a few. For instance, in a wide-ranging review of why the rates of childhood malnutrition in south Asia are twice what they are in sub-Saharan Africa, Ramalingaswami et al. (1997) conclude that “the exceptionally high rates of malnutrition in South Asia are rooted deep in the soil of inequality between men and women” (p.16).

If programs that aim to increase women’s income today can increase women’s status within the household, then the returns to those programs tomorrow could be profound in terms of lowered fertility, improved child survival, and increased human capital.

References


