Identifying Motherhood and its Effect on Female Labour Force Participation in South Africa: An Analysis of Survey Data

by

Gabrielle van der Stoep

School of Economics and Finance
The University of KwaZulu-Natal, Durban
December 2008

Declaration

This thesis is submitted in accordance with the requirements of a Masters of Commerce (by thesis) in Economics at the University of KwaZulu-Natal, Durban. I declare that this is my own work, except where acknowledged in the text, and that this thesis has not been submitted for a degree at any other university.

Gabrielle van der Stoep
11 December 2008

Contents

List of Tables

List of Figures

Acronyms

ACDIS	Africa Centre Demographic Information System
AHDSS	Agincourt Health and Demographic Surveillance System
CAPS	Cape Area Panel Study
CEB	Children ever born
GHS	General Household Survey
IV	Instrumental Variable
KIDS	KwaZulu-Natal Income Dynamics Study
LFS	Labour Force Survey
NIDS	National Income Dynamics Study
OECD	Organisation for Economic Co-operation and Development
OHS	October Household Survey
OLS	Ordinary Least Squares
PSLSD	Project for Statistics on Living Standards and Development
SALDRU	The South African Labour and Development Research Unit

Acknowledgements

I would like to acknowledge key individuals who have provided support and guidance to me over the course of my Masters thesis. I am indebted to Professor Dorrit Posel whose dedication and commitment has exceeded her supervisory role. She has fundamentally guided my often haphazard ideas and has provided rigorous and constructive commentary throughout the course of my work. I am especially grateful for the opportunity to have written a paper with her during this year which resulted in wonderful exposure to both local and international conferences. This was a pivotal point in my academic studies which has immensely improved my analytical ability, presentation skills and academic writing skills. Professor Dorrit Posel has been extraordinarily patient in allowing me to learn from my mistakes; and her generosity in various ways has made the past two years a memorable experience. I would also like to thank her as well as Daniela Casale, Colette Muller, Claire Vermaak, Beverly Edkins and Jim Fairburn for allowing me to attend an informal econometric study session. Learning together facilitated my understanding of instrumental variable regression and selection issues in econometric analysis.

My parents have played an important role in encouraging my studies and praising excellence throughout my university career. I also acknowledge Bradford Wills, my beloved fiancé, who has cheered me along whether a new page was complete or even thrown away. Thank your continued prayers, space to work and for enthusiastically understanding my thesis subject. You have also played a vital role in helping me develop and maintain a work timetable. Finally, thank you to my Father in heaven whose promises in 2 Chronicles 20 hold true.

This thesis was made possible by the generous National Income Dynamics Study (NIDS) scholarship for Masters Study over the period 2007 and 2008. Additional funding was made available by the DAAD scholarship programme for Masters study in 2007 and 2008, a National Research Foundation grant holder scholarship and the University of KwaZulu-Natal Masters Scholarship Award.

Abstract

The objective of this thesis is to investigate the relationship between motherhood and women's labour force participation in South Africa. The key problem in estimating this relationship is the endogeneity of motherhood/childbearing with respect to women's labour force participation. Childbearing behaviour and decisions to participate in the labour force are jointly determined; and unobservable characteristics which influence childbearing behaviour are also correlated with women's labour force participation. This thesis shows that the definition of motherhood can exacerbate these sources of endogeneity bias. International studies typically identify mothers as women with biological children aged 18 years or younger who are co-resident with at least one of their children. In South Africa, however, a sizeable sample of women is not co-resident with their children. The remaining sample of co-resident mothers are a non-random sample of all mothers who are less likely to participate in the labour force than all mothers. Placing a co-residency restriction on motherhood therefore biases the relationship between motherhood/childbearing and labour force participation. In particular, it overestimates the negative relationship. In the international literature instrumental variable (IV) estimation has been used to disentangle these causal mechanisms. This thesis also considers an application of same sex sibling composition, first introduced by Angrist and Evans (1998), as a strategy to identify the exogenous effects of childbearing on women's labour force participation in South Africa.

Little or no research has investigated this relationship in South Africa. One possible explanation for this is that studies on female labour force participation in South Africa have not been able to match women to their children with the datasets that have been analysed: most nationally representative household surveys in South Africa do not contain detailed birth history information. The first part of this thesis analyses what data are available to identify women with children and the quality of these data; it also outlines four different methods to match women to their children using these data. The second part of this thesis investigates the relationship between motherhood/childbearing and women's labour force participation in South Africa.

Introduction

Since the 1990's a number of nationally representative household surveys have been made available to analyse South African labour market activity. Accordingly, there has been a growth in literature on female labour force participation in South Africa. Studies have investigated the key determinants of female labour force participation and explored reasons for its dramatic rise post-apartheid (see for example Casale 2003; Bhorat and Leibbrandt 2001; Winter 1999). Despite the feminisation of the South African labour force, labour force participation rates of men still exceed those of women (Casale 2004). One possible reason for this is that women disproportionately bear the responsibilities of bearing and raising children which reduces opportunities to work or to search for work. As its main objective, this thesis investigates the effects of motherhood (or childbearing) on female labour force participation in South Africa using the General Household Survey 2002.

In general, a well established literature finds a negative relationship between childbearing/fertility and labour force participation, especially among women with very young children; however the interpretation of this relationship is complicated by the endogeneity of childbearing behaviour. Economists and demographers have acknowledged that a simultaneous causal relationship exists between childbearing behaviour and decisions to participate in the labour force. As a result of this two-way relationship, childbearing is correlated with the error term in a linear equation of female labour force participation. Ordinary least squares estimation therefore leads to biased estimates of the effects of childbearing on women's decisions to participate in the labour force. Furthermore, heterogeneity in unobserved characteristics across different groups of women further biases the estimation of the true exogenous effects of childbearing on labour force participation.

A key research question in this thesis asks how the definition of motherhood itself can exacerbate these sources of endogeneity bias. International studies typically place a co-residency restriction on definitions of motherhood used in female labour force participation estimations. There are both logical and practical reasons for this restriction.

First, children are not expected to constrain the labour force activity of mothers who are not living with their children. Second, co-residency with children may also be a necessary requirement to match women to their children with available datasets. Many household datasets do not contain fertility information. Instead information on relationships among resident household members is used to identify mothers (see for example Vere 2008; Angrist and Evans 1998; Bronars and Grogger 1994); but identifying mothers in this way requires that mothers are resident with their children in the same household. In South Africa, however, a history of labour migration and other socio-economic factors have served to fragment family living arrangements, in particular the co-residency of mothers and children. As a result a sizeable sample of all mothers in South Africa is not co-resident with their children. In 2002, about 15 percent of all African mothers aged 20 to 49 were not co-resident with any of their children aged 18 years or younger. This thesis shows that the remaining sample of co-resident mothers are a non-random sample of all mothers who are significantly less likely to be labour force participants than non-mothers *and* not co-resident mothers. A co-residency requirement therefore biases the relationship between motherhood and labour force participation.

In a context of labour migration and extended family structure, fosterage of children by other women, including non-biological mothers, has been a common occurrence in South Africa. I investigate the implications of fostering children (or surrogate parenthood) on women's labour force participation in South Africa and show that non-mothers who are living with children are significantly less likely to be labour force participants than non-mothers who do not live with any children.

In addition to showing how the definition of motherhood can exacerbate sources of endogeneity bias, this thesis uses instrumental variable (IV) estimation to explore the exogenous effects of childbearing on labour force participation. First introduced by Angrist and Evans (1998), I explore whether the sibling sex composition of a woman's first two children is an appropriate strategy in the South African context to identify the exogenous effects of childbearing in female labour force estimations.

The thesis is structured as follows. Chapter One provides a review of both theoretical and empirical literature on childbearing and female labour force participation. The review highlights the presence of a two-way causal relationship between childbearing behaviour and decisions to participate in the labour force which complicates the estimation of this relationship. A key focus of the chapter is to review international studies which have attempted to disentangle these causal mechanisms using instrumental variable estimation. Chapter One also highlights that scant attention has been paid to this relationship in South Africa, largely due to data constraints. Studies on female labour force participation in South Africa typically have not been able to match women to their children with datasets they have used. In direct response to this problem, Chapter Two analyses what data are available to identify mothers in South Africa and the quality of these data. I describe two sets of methods that can be used to identify mothers as well as their associated benefits and limitations. The first set uses birth module/fertility information to identify biological mothers, regardless of co-residency with their children. Consistent with international literature, the second set uses questions on household relationships to identify mothers who are co-resident with their children.

Chapter Three then describes motherhood status in South Africa using the different identification methods. A key contribution of this chapter is to show that measures of mothers are very sensitive to which of the identification methods is used. In particular, the second set of methods, which restrict mothers to women who are co-resident with their children, substantially underestimates the true number of biological mothers in South Africa.

Chapter Four proceeds to compare the characteristics of non-mothers, co-resident and not co-resident mothers, highlighting the distinctive characteristics of mothers who are co-resident with their children. In particular, it describes differences in labour force participation rates across these different groups of women. This descriptive analysis provides initial evidence that childbearing is associated with opportunity costs including forgone education and lower labour force participation rates. It also explores the living arrangements of women with children who are not their own. I identify if residing with

children who are not one's own is correlated with lower labour force participation rates among non-mothers.

In Chapter Five I use econometric estimations to estimate the relationship between motherhood status and labour force participation in a more rigorous multivariate framework. I investigate not only how childbearing lowers the probability of labour force participation among mothers, but also among non-mothers who live with children. The key contribution in Chapter Five is to show how the definition of motherhood can make a particularly big difference to estimates of female labour force participation. However the econometric analysis does not directly control for the endogeneity of childbearing in labour force estimations. In Chapter Six, I extend the analysis and evaluate whether IV estimation can be used to generate unbiased estimates of the motherhood effect.

Chapter One:

Literature review

Introduction

The relationship between childbearing (or fertility) and labour force participation has been of longstanding interest in the social sciences. Labour economists have typically treated childbearing as a key determinant of whether women participate in the labour force and how many hours of work they choose to supply, while demographers have treated women's labour force participation as a key determinant of their childbearing behaviour (Browning 1992; Cramer 1979). In the past three decades, however, a synthesis of literature has developed which acknowledges the causal interdependence of childbearing behaviour and labour force decisions. In the empirical literature, in particular, much focus has been given to disentangling this causal mechanism in order to estimate an exogenous effect of childbearing on women's labour force participation.

The key objective of this chapter is to review both the theoretical and empirical literature on this relationship. I first review the economic theory of labour supply in section 1.1 with the view to understanding how childbearing influences a woman's decision to participate in the labour market. I describe the traditional neoclassical theory of labour supply, followed by Becker's theory of the allocation of time and joint labour supply decisions within the household.

In section 1.2, I review empirical literature which estimates the relationship between childbearing and women's labour force participation. A large number of empirical studies have investigated this relationship. These studies can be distinguished into three types: those that investigate this relationship at the time series level; the cross-national level and at the individual level. Time series analyses typically investigate how trends in fertility have explained trends in female labour force participation, in particular the rise in female labour force participation since World War II (see for example Doepke *et al*

2007; Leibowitz and Klerman 1995). Cross-national studies have explored how varying associations between rates of fertility and rates of labour force participation can be explained by cross-country variations in family policy, child cash benefits and childcare arrangements (see for example Del Boca and Locatelli 2006; Brewster and Rindfuss 2000; Lloyd 1991); and studies at the cross-section, or individual level, focus on the causal relationship between childbearing/fertility behaviour and a woman's decision to participate in the labour force (see for example, Aguero and Marks 2008; Boushey 2008; Posel and van der Stoep 2008; Vere 2008; Cruces and Galiani 2007; Connelly *et al* 2006; Chapman *et al* 2001; Angrist and Evans 1998; Gormick *et al* 1996; Bronars and Grogger 1994; Cain and Dooley 1976; Mincer 1962). In section 1.2, I focus specifically on studies that investigate this relationship at the individual level in both developed and developing country contexts. Attention is given to cross-sectional studies that have addressed the endogeneity of childbearing status in estimations of female labour force participation.

Following the review of the international literature, section 1.3 considers specifically women's labour force participation in South Africa and the context in which it has occurred. I identify gaps in the South African literature on the effects of childbearing on women's labour force participation and highlight key aspects of the South African context which may influence a study of this relationship.

1.1 Review of theoretical literature

1.1.1 Theories of labour supply: understanding the effects of children on labour supply decisions.

Labour supply decisions can be divided roughly into two categories: the decision about whether to participate in the labour force at all; and if so, how many hours of labour to supply. Theories of labour supply generally focus on the decision about how many hours of labour to supply to the market; however, they are also useful for explaining which individuals participate in the labour force and what factors influence their participation

decisions (Casale 2003:7). I present two models of labour supply, the traditional neoclassical model of individual labour supply followed by the model of household production. I consider specifically how each theory facilitates an understanding of the effects of children on female labour force participation.

The decision about whether or not to participate in the labour force is ultimately a decision about how to spend time. In the traditional neoclassical model of labour supply the decision to work involves the optimal allocation of time across two broad activities: work and leisure. Leisure is an all-encompassing term that includes every activity other than supplying labour to the market and is assumed to provide a direct source of satisfaction or utility to the individual. Subsumed within leisure is childcare which is also assumed to provide direct utility to the individual. Time spent working provides an indirect source of utility as the earned wage is used to purchase consumption goods which provide utility. In this theory, the individual is the decision-making unit who will aim to maximise utility derived from both leisure and work, subject to a scare resource of time. The decision to work and how many hours to work is therefore presented as a constrained utility maximisation problem (Cahuc and Zylberberg 2004; Reynolds *et al* 1998; Ehrenberg and Smith 1994).

An individual cannot pursue both leisure and market work at the same time; each hour engaged in one of the two activities will result in a costly trade off of the other. The opportunity cost of each additional hour of leisure time is the hourly wage rate which could have been earned; while the opportunity cost of an additional hour of work is the satisfaction lost from forgone leisure time. The optimal allocation of hours between leisure and work, that maximises benefits while minimising costs, will be determined by the prevailing market wage rate and the availability of non-labour income (Cahuc and Zylberberg 2004; Casale 2003; Reynolds *et al* 1998). Holding income constant, a rise in the wage rate will increase the opportunity cost of leisure so that the individual will substitute hours of work for leisure, increasing his or her labour supply. This is called the *substitution effect* which positively affects hours of labour supplied. However an increase in the wage rate will also increase an individual's income, enabling the individual to purchase more leisure time. This is termed the *income effect* and has a

negative effect on hours of labour supplied. The overall impact of a rise in the wage rate on hours of labour supplied will depend on which of the two effects dominates. If the substitution effect is greater (smaller) than the income effect, hours supplied to the labour market will increase (decrease) and less (more) time will be devoted to leisure (Ehrenberg and Smith 1994:169-187).

Although the basic work-leisure model is used to determine how many hours of work will be supplied, it can also be used to determine if an individual will participate in the labour force (i.e. supply more than zero hours of work or no hours of work). An individual will not consider working unless the market wage rate offered exceeds an individual specific reservation wage - the lowest hourly wage at which an individual is willing to supply his or her first unit of labour to the market (Ehrenberg and Smith 1994:180; Gronau 1973). If two individuals face the same potential wage rate in the labour market, the individual with the higher reservation wage will be less likely to enter the labour market. Compared to women without children, women with children typically face a higher reservation wage. Monetary costs of formal childcare and time costs associated with taking children to childcare facilities raise the cost of working and therefore the reservation wage (Ehrenberg and Smith 1994:206-211). Holding all else constant, neoclassical labour supply theory predicts that women with children will have a lower probability of participating in the labour market than women with no children.

One of the main criticisms of the traditional neoclassical model of labour supply is that it fails to interrogate the different ways in which time is spent outside work. Non-market time involves two distinct activities: household consumption which includes sleeping and entertainment; as well as production such as cleaning, cooking and childcare. The latter is sometimes termed *unpaid work* for two reasons: first, it requires energy or effort (Cahuc and Zylberberg 2004) and second, it satisfies the 'third person criteria' where this work could be outsourced to someone else (Gronau and Hamermesh 2001). By contrast, household consumption activities typically do not require much effort and cannot be performed by a third person. Despite this distinction, household production activities are subsumed within the neoclassical notion of "leisure". Consequently, the

neoclassical model of labour supply assumes similar trade offs between market work and household consumption (or pure leisure) as between market work and household production. In reality, pure leisure and market work are not as highly substitutable as household production and market work. Household production can be reduced, for example, by hiring domestic help such as nannies or au pairs; or by purchasing ready cooked meals and labour saving technologies such as washing machines (Casale 2003:13; Reynolds 1998:57).

Another major criticism of the neoclassical model of labour supply is that the allocation of time is not a decision made by individuals independently of the decisions of other individuals in the household. In 1965, however, Gary Becker's seminal theory on household production and the theory of the allocation of time addressed these limitations of the traditional neoclassical model of labour supply. The model of household production explicitly recognises that both consumption and production take place in the household. Here time is allocated across three basic activities - market work, household production and pure leisure - and the family rather than individual is the sole agent that maximises its own total utility (Cahuc and Zylberberg 2004:15). The family's total utility is a function of the consumption of leisure and basic 'commodities' that are produced in the household by combining time with purchased market goods. The family will optimally allocate each household member's time across the three activities to maximise its total utility function subject to a family wealth constraint. Decisions about which members engage in market work or household work, and how many hours are devoted to each activity, will depend particularly on the relative productivities of family members in market work or household production (Ehrenberg and Smith 1994:212-223; Becker 1965:512). Women obviously have a comparative advantage in childbearing but have also developed a comparative advantage in other aspects of household production as a result of socialisation, preferences or labour market disadvantages. Furthermore, men have historically faced higher earnings opportunities, perhaps as a result of discrimination against women in the market place (Reynolds 1998:57). The model of household production therefore predicts a sexual division of labour across couples:

women allocate most of their time to non-market work while their husbands/partners allocate their time to market work.

In Becker's model of household production a principal assumption is that a unified household utility function can be specified and family members pool their resources of time and income into one family budget. This is possible where individual family members face identical preferences. However where family members have individual preferences, Becker argues that the household is still able to maximise this joint utility function as each household member conforms to the preferences of an altruistic household head or 'benevolent dictator'. An optimal allocation of resources is assumed to occur as the benevolent dictator transfers resources so as to induce everyone to act in the best interests of the household (Casale 2003:23; Schultz 1990:600). Many studies, however, contest this view of how the household works and argue that an aggregated family utility function may be unjustified.[1] While household members and the household head may behave altruistically, they may also direct household transfers to further self interests which conflict with overall household interests. For example, as a means of exerting power over women in household decision-making, husbands may prevent their wives from engaging in market work. In this case, the sexual division of labour is determined by reasons other than relative productivities in market and non-market work (Casale 2003:24).

Despite limitations of Becker's model of household production, it still provides a better theoretical framework within which to incorporate fertility behaviour into female labour supply decisions than the traditional neoclassical theory of labour supply (Willis 1973). By recognising that non-market time involves time in household production such as childcare, rather than just leisure time, the effects of childrearing on labour supply decisions can be considered more carefully. Children yield satisfaction (or otherwise termed 'psychic income') to their parents (Gronau 1973; Willis 1973; Becker 1960) and

[1] In response, bargaining models have been developed which acknowledge that self interest may exist in family resource allocation. These models use Nash-equilibrium game theory to understand family decision-making about labour supply (see for example Schultz 1990).

can be entered as a variable in a household utility function. This satisfaction can be increased by having more children and supplying more resources, namely time and goods, to produce better quality children or otherwise termed child 'services' (Willis 1973:S14-S25). Unlike the traditional neoclassical model of labour supply, children are entered as a separate variable in a household utility function, rather than being subsumed within a general 'leisure' variable. Therefore explicit trade-offs between childcare (or the production of "child services") and labour supply can be modelled.

A woman's time in non-market work is particularly valuable to the household for producing utility-deriving child services. However, the amount of time a woman allocates across home and market work will depend critically upon the price or value of her time in non-market work (Willis 1973:S31; Becker 1965) which is a positive function of the amount of child services produced in the household. The value or price of time is analogous to the term 'reservation wage' used in the basic work-leisure model. A necessary condition for a woman's entry into the labour force (i.e. labour force participation) is that her potential wage exceeds her price of time in non-market work (Gronau 1973; Willis 1973:S33). But the price of time will increase above the wage rate if the family chooses large amounts of child services i.e. many children and high 'quality' children. Holding all else constant, the likelihood that a woman participates in the labour market will fall with the production of increasing amounts of child services.

The price of a woman's time in non-market activities will also depend on the age composition of her children (Gronau 1973:S183). Where time and market goods are key inputs in the production of child services, their substitutability will vary with the age of the child. When children are very young, the elasticity of substitution between inputs of time and market goods is very low because young children are particularly time intensive rather than goods intensive. As children grow older, inputs of time can be more easily substituted with market goods in the production of child services (Browning 1992:1454). Younger children are therefore expected to have larger positive effects in raising the price of a mother's time (or reservation wage) above the prevailing market wage and lowering her probability of participating in the labour force (Gronau

1973:S188). As children grow older they may have no positive effect on the price of her time in non-market activities. Women become more experienced or learned in the production of child services as children get older. Their productivity in child services rises with 'on-the-job' training, allowing them to produce the same amount of child services with decreasing amounts of time inputs (*ibid*:S188). As children grow older it also becomes easier to substitute paid childcare for unpaid childcare. Furthermore children themselves may also begin to contribute to home production with increasing years of schooling, offsetting the demand for a mother's time in non-market activities.

Becker (1985) also suggests that the probability of labour force participation will be lowest among mothers with very young children because time spent with infants exhausts a woman's energy. Little energy is therefore reserved for market work and productivity is compromised on the job. Mothers may obtain alternative childcare, but the intensive supervision required by young children increases formal childcare costs which discourages women from working (Connelly 1992; Browning 1992:1458). Browning (1992:1458), however, notes that even if childcare costs were constant across the age of children, mothers may still prefer to stay at home with their children when they are younger. 'Psychic income' or utility derived from time spent with children may be higher when children are younger (*ibid*:1458; Gronau 1973:S188). For these reasons, the probability that a mother participates in the labour market will be lower the younger the age of her child.

1.1.2 The endogeneity of childbearing/fertility and labour supply

Both the traditional neoclassical theory of labour supply and the household production model predict that the presence of children will be negatively related to women's labour force participation as it raises a woman's reservation wage or price of non-market time; and this relationship is expected to be particularly strong among mothers with very young children. Further developments in both the economic and demographic literature, however, extend our understanding to recognise that the decision to have children itself may be a function of women's labour force participation. In fact demographers often

explicitly take female labour supply behaviour as a key determinant in fertility models (Browning 1992:1451). The negative effect of female labour supply on fertility is explained by the role incompatibility hypothesis (Stycos and Weller 1967). Mason and Palan (1981:551) note that

> "this hypothesis posits that an inverse relationship between women's work and fertility occurs only when the roles of worker and mother conflict, this being a situation in which women are forced to make trade-offs between their participation in productive employment and the number of children they bear".

The degree to which the roles of worker and mother conflict will depend on the nature of work and the social organisation of child care arrangements (Stycos and Weller 1967:215). Where worker and mother roles are highly incompatible, labour force participation will depress fertility. One possible mechanism by which this occurs is that working women may delay the transition into motherhood (Brewster and Rindfuss 2000:282).

In the economic literature, labour force participation may raise the costs of childbearing through its effect on wages, or work may reduce the time and energy available for childbearing (Becker 1985, 1960; Willis 1973). Recognising that childbearing/fertility may be both a cause and effect of labour force participation, economists have also modelled fertility behaviour and labour supply decisions as causally interdependent (see for example, Willis 1973). However this complicates the empirical estimation of this relationship. Consider the following ordinary least squares (OLS) equation for women's labour force participation:

$$Y_i = \alpha + \beta M_i + \delta X_i + e_i \qquad (1)$$

Y_i is a dependent variable taking on a value of one if a woman is labour force participant and zero if she is economically inactive. M_i reflects whether or not a woman has a child, the number of children she has or the age of her children. X_i is a vector of other

exogenous variables and e_i is a random error term. The coefficient, β, is the motherhood effect (or child effect); it captures the effect of childbearing on women's labour force participation. Because childbearing decisions and labour force participation decisions are simultaneously determined, childbearing variables, M_i will be correlated with the error term e_i in equation (1). OLS estimates of β will therefore be biased and inconsistent as a result of the two-way causal relationship between fertility behaviour and labour force participation decisions (Wooldridge 2003). Endogeneity in childbearing may also arise due to unobserved heterogeneities across different groups of individuals (Aguero and Marks 2008; Browning 1992:1451). Becker (1985:S49) states that

"the stock of energy varies enormously from person to person, not only in dimensions like mental and physical energy, but also in 'ambition' and 'motivation'."

For example mothers may have less ambition, motivation or energy than non-mothers (Aguero and Marks 2008; Posel and van der Stoep 2008) which will lower their probability of labour force participation. If the effects of unobserved heterogeneity are not controlled for in labour force estimations, then estimations of the direct effect of childbearing are confounded by spurious factors. In this case, ordinary least squares (OLS) or probit estimations will overestimate the negative effect of motherhood or childbearing on women's labour force participation. This discussion is elaborated in more detail in Chapter Six.

1.2 Review of empirical literature on childbearing and female labour force participation

A large number of studies have estimated the relationship between motherhood/ childbearing and labour force participation in both developed and developing countries.[2]

[2] In addition to studies that investigate the effects of childbearing on labour force participation, there is an extensive literature which estimates the opportunity cost of motherhood in terms of forgone earnings (see for example Piras and Ripani 2005; Anderson *et al* 2003; Anderson *et al* 2002; Joshi 2002; Budig and England 2001; Chapman *et al* 2001; Waldfogel 1997; Korenman and Neumark 1992; Cramer 1979).

In the following subsections I review this international literature and give particular attention to studies that have attempted to disentangle the causal mechanisms between childbearing and labour force participation in section 1.2.3.

An important clarification to make is that the terms labour force participation and employment are generally used interchangeably in the international literature considered below. Employment may refer only to work for pay (Cruces and Galiani 2007; Angrist and Evans 1998) or may be extended to include not working for pay, for example in a family farm or business (Wong and Levine 1992). Labour force participation includes both employment and unemployment. A possible reason for treating the terms employment and labour force participation as synonymous is that unemployment rates are typically low, particularly in developed countries. In South Africa, however, a significant proportion of labour force participants are unemployed, where unemployment rates are highest among women (Casale and Posel 2005). A decision to participate in the South African labour force, therefore, does not automatically imply that an individual will be employed.

1.2.1 Childbearing and female labour force participation in developed countries

Consistent with the economic theory of labour supply, there is conclusive evidence of a significant negative relationship between motherhood or childbearing and women's labour force participation in developed countries, controlling for a range of covariates (Boushey 2008; Chapman *et al* 2001; Angrist and Evans 1998; Gormick *et al* 1996; Bronars and Grogger 1994; Lloyd 1991). In general, studies show that this negative relationship is robust to whether motherhood/childbearing is treated as exogenous or endogenous with respect to labour force participation. It is also influenced by the age and number of children born to a woman and is found to be particularly strong if women have very young children (Boushey 2008; Chapman *et al* 2001; Gormick *et al* 1996; Bronars and Grogger 1994; Rosenzweig and Wolpin 1980).

Boushey (2008), for example, estimates the effect of motherhood on the probability of employment among women in the United Sates from 1979 to 2005. She controls for race, age, marital status, the log of other household income, the number of non-working adults in the household, education and region in labour force participation estimations. Her study, however, does not control for possible endogeneity of childbearing. She finds a negative marginal effect of having at least one child under the age of 18 during the period of analysis (*ibid*:9); but the size of the effect is considerably larger if a woman had a child under the age of six.

Interestingly, Boushey (2008) also finds that the motherhood effect has fallen over time. Having children under the age of 18 reduced the probability of employment by about 22 percent for the average woman in 1979 but this declined to less than 14 percent by the late nineties (*ibid*:1). These results are consistent with research in other OECD countries which identifies that the strong negative correlation between childbearing and labour force participation is growing weaker with time (Del Boca and Locatelli 2006). This trend has been attributed to changes in the economic constraints faced by women in their labour force participation and fertility decisions (*ibid*:5). For example, the introduction of more child-oriented policies such as parental leave and affordable childcare, have reduced the opportunity costs of childcare. The introduction of more flexible labour market policies (such as part-time jobs) has also helped mitigate conflicts between labour market work and motherhood responsibilities (*ibid*:5).

1.2.2 Childbearing and female labour force participation in developing countries

Although a clear negative relationship exists between childbearing and labour force participation in developed countries, empirical evidence of this relationship is less consistent in developing countries (Bianchi 2000; Lloyd 1991; Mason and Palan 1981). Often no significant relationship is found or at times the relationship is even positive (Aguero and Marks 2008; Posel and van der Stoep 2008; Vere 2008; Connelly *et al* 2006; Bianchi 2000; Wong and Levine 1992; Lloyd 1991; Mason and Palan 1981). Furthermore the strength of this relationship is found to be particularly sensitive to

whether estimations control for possible endogeneity in motherhood status (see for example Aguero and Marks 2008; Connelly *et al* 2006).

In developed countries the negative relationship is generally attributed to an incompatibility or conflict between childcare and labour market activity (Stycos and Weller 1967). In developing countries, however, various socio-economic factors may reduce the tradeoffs between childcare and market work (Lloyd 1991; Mason and Palan 1981). First, the physical separation between labour market activity and household production is less rigid than in developed countries (Posel and van der Stoep 2008). Given the industrial nature of production in developed countries, work sites are situated away from the home and are often not safe for children; rigid work schedules must be followed and the presence of children may jeopardise productivity on the job (Brewster and Rindfuss 2000). Consequently a clear tradeoff exists in the allocation of time across market work and childcare (Mason and Palan 1981:551). But in developing countries, market-related work (such as agricultural production) is often less removed from the household and affords women the opportunity to work in closer proximity to their children (Mason and Palan 1981:551). For this reason market activity may be more easily combined with childrearing.

The reduced conflict between childrearing and market work in developing countries, compared to developed countries, is also attributed to differences in household living arrangements and family structures (Mason and Palan 1981). The co-residence of extended families in developing countries, particularly in Africa, increases the availability of parental surrogates in the household (Posel and van der Stoep 2008; Wong and Levine 1992).[3] In this case, grandparents, older siblings, other relatives or even members of the wider community provide a network of foster or informal childcare, freeing up time for mothers to pursue labour market opportunities (Sudarkasa 2004; Hallman *et al* 2002; Deutsch 1998; Tiefenthaler 1997; Wong and Levine 1992;

[3] One paper that provides a particularly detailed review of the effects of household structure on women's economic activity is Wong and Levine (1992). They analyse the effects of household structure on women's economic activity in urban Mexico.

Lloyd 1991; Mason and Palan 1981). As childcare responsibilities are distributed across household members, the availability of parental surrogates in developing countries may lower the mother's price of time in non-market work, compared to developed countries. This is consistent with household models of labour supply, which predict that the presence of other potential caretakers in the household may lower the amount parents pay for childcare, increasing the probability of their labour force participation (see for example Connelly 1992).

The reduced conflict between working and mothering may also relate to differences in norms or attitudes about appropriate childcare and whether children can contribute to household production (Brewster and Rindfuss 1996; Mason and Palan 1981). For example, the unsupervised play of children or the provision of childcare by older siblings, deemed inadequate childcare in developed countries, may be perfectly acceptable in a developing country context (Sudarkasa 2004; Hallman *et al* 2002; Mason and Palan 1981:569). There is considerable evidence to suggest that a woman's own children provide care-giving roles to younger siblings in developing countries (see for example Bray and Brandt 2007; Hallman *et al* 2002; Mason and Palan 1981). As older children contribute to the household production of child services, it reduces the demands on a mother's time in non-market work, raising the likelihood that she will participate in the labour force.

In developing countries, women typically face lower average household incomes compared to developed countries. They may therefore face a greater economic need to work in order to supplement the income of the household (Hallman *et al* 2002; Deutsch 1998). Mothers will choose not to work only if they can afford to buy non-market time to stay at home caring for children (Deutsch 1998:12).

1.2.3 Controlling for endogeneity in childbearing status

Identifying the strength and causality of the relationship between motherhood/ childbearing and female labour force participation is complicated by the endogeneity of

childbearing (Aguero and Marks 2008; Cruces and Galiani 2007; Angrist and Evans 1998; Cramer 1979). The main approach used to address this endogeneity is instrumental variable (IV) estimation. This involves finding an instrumental variable which satisfies two requirements: first, the instrument must be highly correlated with childbearing but second, it must be uncorrelated with the error term in the labour force participation equation (Wooldridge 2003). Instruments used to identify an exogenous effect of childbearing on labour force participation have included twin births (Connelly *et al* 2006; Bronars and Grogger 1994), same sex sibling composition (Cruces and Galiani 2007; Iacovou 2001; Angrist and Evans 1998), fecundity or infertility status (Aguero and Marks 2008) and most recently the Chinese Lunar calendar (Vere 2008). In the analysis below, I review studies using these different instruments. Where studies have estimated both OLS and IV estimates, I highlight how a negative relationship between childbearing and labour force participation is overestimated in the presence of endogeneity. I also consider the limitations and benefits of each of the various instruments.

Introduced by Rosenzweig and Wolpin (1980), twin births was one of the first instruments used to explore the exogenous effects of childbearing on women's labour force participation. Twin births are argued to be truly exogenous with respect to a woman's unobserved characteristics[4] yet they have a strong positive effect on the number of children born to a woman (Bronars and Grogger 1994). Bronars and Grogger (1994) use twins at first birth to determine the effect of an unplanned (exogenous) birth on the economic outcomes of young single mothers in the United States. They find a significant negative *exogenous* effect of a child aged three years on the probability that a mother works or looks for work, but by the time the child reaches ten to 13 years old the unplanned birth has no effect on the mother's participation (*ibid*:1152). More recently, Connelly *et al* (2006) use twins at first birth to identify the effects of recent fertility on women's employment in Brazil. They find that IV estimates of the effects of recent

[4] The probability of having twin births increases with a woman's age (Bronars and Grogger 1994). Twin births are therefore only truly exogenous with respect to unobserved characteristics if labour force estimations control for age at first birth.

childbearing on women's employment are small and insignificant. By contrast if childbearing is treated as exogenous using OLS estimations, they find a large negative relationship between recent fertility and women's employment in Brazil (*ibid*:583).

The sibling sex composition of children has also been exploited as an instrument for further childbearing among women. It is argued that parents whose first two children are of the same sex are more likely to have another child than parents with a mixed sibling sex composition because parents have a 'taste' for a balanced or mixed sibling sex composition (Ben-Porath and Welch 1976). If parents have a preference for a mixed sibling sex composition but only have a same sex sibling composition, the total desired utility from having a certain number of children is lowered. This raises the marginal utility of further childbearing, increasing the probability that additional children will be born (Angrist and Evans 1998:456). Thus same sex sibling composition is positively correlated with childbearing. It is also uncorrelated with unobserved characteristics because of the biological random assignment of children's sex composition. Using same sex sibling composition as an IV, Angrist and Evans (1998) find that in the United States having more than two children yields a negative effect on whether or not a woman participates in the labour force; but the IV estimate is less negative than the OLS estimate.[5] Same sex sibling composition has also been used to estimate the effects of further childbearing on women's labour force participation (or employment) in the United Kindgom (see Iacovou 2001), Argentina and Mexico (see Cruces and Galiani 2007). Consistent with the United States, IV estimates using same-sex sibling composition are negative but smaller in magnitude than OLS estimates in these countries.

[5] Angrist and Evans (1998) compare IV estimates using same sex siblings with twins at second birth. Both instruments measure the consequence of moving from two to three children but they find that estimates using twins are consistently less negative than estimates using same sex siblings. Differences in estimates may relate to economies of scale in parenting children of the same age or because a third child, born as a consequence of multiple births, is older than a third child born to families not experiencing a multiple birth (Angrist and Evans 1998:471-472).

These studies using twin births or same sex sibling composition have provided useful evidence that treating childbearing as exogenous biases upward the negative relationship between childbearing and labour force participation. However application of these IV's suffers from several limitations. First, the occurrence of multiple births is a rare phenomenon; as a result the application of twin births as an instrument for childbearing requires very large datasets which often are not available in many countries (Connelly *et al* 2006; Browning 1992), including South Africa. Using same sex sibling composition relies on the assumption that a strong preference for a mixed sibling sex composition exists among women or couples. However this assumption may not hold in all countries (Ben-Porath and Welch 1976). For example, in many African countries with high average levels of desired fertility, parents may be unconcerned about whether the first two children are of the same sex because eventually they are likely to have a child of the opposite sex (Arnold 1992). I evaluate this argument in further detail in Chapter Six.

The most severe limitation of using same sex siblings or twin births as instrumental variables for childbearing, is that they can only be used to instrument for *further* childbearing among women with *at least* one or two children (one child in the case of using twins at first birth, and two children in the case of using same sex siblings or twins at second birth). However sibling sex composition and the presence of twins are unrelated to whether a woman shifts her fertility behaviour from zero children to one child; therefore they are not suitable instruments for determining the effects of having any children, as opposed to no children at all, on women's labour force participation (Aguero and Marks 2008; Vere 2008:304, 312). This has implications for estimating a motherhood effect because the impact of higher order births on women's labour force participation may differ considerably from that of a first birth. Vere (2008:304-5) summarises the argument as follows:

> "The gap in the labour supply literature is that, although the sex mix instrument can identify the causal effect of a third- or higher order birth, and the twins instrument can identify the effect of birth of second order or higher, inference about the effect of the first birth is impossible with these

> instruments. IV estimates of the marginal effects of higher order births are still conditional on a woman's decision to have a first child. The marginal effect of the first birth, therefore, is of critical importance in determining the total effect of fertility on female labour supply. Moreover, if there are economies of scale to be realised in raising children, the marginal effect of the first birth may be much higher than the effects of subsequent births."

Very recent studies in developing countries have identified IV strategies to determine the marginal effect of the first birth (or whether or not a woman has any children at all). Employing a unique IV strategy that uses the Chinese Lunar calendar, Vere (2008) is able to identify the marginal effect of a first birth on women's labour force participation in Hong Kong. According to Chinese astrology, children born in certain years, in particular the year of the dragon, are considered to have superior characteristics or qualities. Therefore certain years are perceived as being more auspicious than other years for childbearing. Using a Cox proportional hazard model, Vere (2008) finds that Chinese women in Hong Kong are more likely to give birth in dragon years than in non-dragon years. Thus Chinese ethnicity becomes a valid strategy for identifying the exogenous effect of childbearing on labour force participation. A negative relationship is found between childbearing and labour force participation using this strategy; but the effect is stronger for the first birth than for subsequent births (*ibid*). This is consistent with earlier studies, for example Cramer (1979), which find that the time costs of children diminish with each successive birth indicating the presence of economics of scale in raising children.

An obvious limitation of Vere's (2008) identification strategy is that its application is limited to countries with dense Chinese populations. A more universally applicable IV strategy, to identify women with at least one child, is suggested by Aguero and Marks (2008). They propose that the infertility status of women is an appropriate instrument for whether or not a woman is a mother i.e. has one or more children. A woman's infertility is obviously negatively correlated with childbearing. It is also arguably uncorrelated with unobserved factors captured in an error term that influence labour force

participation because infertility is a random occurrence (*ibid*). Using infertility to instrument for motherhood, Aguero and Marks (2008) find no causal relationship between motherhood/childbearing and women's labour force participation in six Latin American countries. This result holds regardless of the number or age of children born to women (*ibid*). By contrast, OLS regression results consistently yield a significant negative coefficient on motherhood or child variables, regardless of the number or age of children. These results suggest that the negative relationship between motherhood and labour force participation, observed in studies that do not instrument for motherhood, may be *entirely* the result of spurious factors.

1.3 Female labour force participation in South Africa

The above discussion has highlighted the vast and growing number of studies on the relationship between motherhood and labour force participation in both developed and developing countries. In particular, I have reviewed different identifying strategies to estimate the exogenous effects of childbearing on labour force participation using IV estimation. A natural question to ask at this point is whether studies have investigated this relationship in South Africa? Furthermore, what is the specific context in which female labour force participation has occurred in South Africa? I address these questions in this section.

1.3.1 Empirical studies on female labour force participation in South Africa

Despite a growing literature on female labour force participation in South Africa, scant attention has been paid to the effects of motherhood or childbearing. Studies have typically controlled for the presence of *any* children in the household in female labour force estimations (Casale 2003; Bhorat and Leibbrandt 2001; Winter 1999); but they have not identified the effect of a woman's own biological children on her labour force participation. A possible explanation for the dearth of economic literature on motherhood is that studies have not been able to match women to their own children

with the datasets that have been analysed: most nationally representative household surveys in South Africa do not contain detailed birth history information.

One exception is the study by Mlantsheni and Leibbrandt (2001) which explicitly investigates the economic effects of childbearing on the labour force activity of African women in South Africa using the October Household Survey (OHS) 1995. Unlike international studies in which employment and labour force participation are used interchangeably, Mlantsheni and Leibbrandt (2001) treat labour force participation and employment as two separate processes. Given high rates of involuntary unemployment in South Africa, a decision to participate in the labour force does not automatically imply that an individual will be employed (Bhorat and Leibbrandt 2001; Mlantsheni and Leibbrandt 2001). Mlantsheni and Leibbrandt (2001) therefore estimate the effects of having a young child on a woman's probability of employment using a maximum likelihood probit model but controlling for selection into labour force participation. The results of their selection equation estimation show no significant negative relationship between having children under the age of six and women's labour force participation. The authors suggest that this indicates that African women with children cannot afford to stay at home and care for children. Rather the need to support children encourages mothers to participate in the labour force. This would be consistent with the findings of Deutsch (1998) and Hallman *et al* (2002) who argue that not working is a luxury good for many mothers in developing country contexts.

The limitation of Mlantsheni and Leibbrandt's study is that they only consider the effects of young children, rather than all children, on women's labour force participation. They therefore do not estimate a motherhood effect in South Africa. Their study also uses the OHS 1995 in which the quality of birth data is highly questionable and may compromise the identification of true child effects (see Chapter Two and Three for an analysis of birth data quality in the OHS 1995). Most importantly they neither address the endogeneity of childbearing decisions in labour force estimations, nor suggest possible ways in which to control for endogeneity given available data in South

Africa. A substantial gap therefore remains in the literature on the relationship between childbearing and labour force participation in South Africa.

1.3.2 Childrearing and female labour force participation in the South African context

The international economic literature has typically placed a co-residency restriction on motherhood in labour force participation estimations (see for example Boushey 2008; Cruces and Galiani 2007; Angrist and Evans 1998). One possible reason for this is that questions on household relationships have typically been used to match women to their children (see for example Vere 2008; Cruces and Galiani 2007; Angrist and Evans 1998). A necessary requirement to identify mothers in this case is that children are co-resident in the same household as the mother. Another obvious reason for this restriction is that childbearing is not expected to constrain the labour force participation of mother's who are not co-resident with their children because they are not involved in day-to-day childrearing activities. In South Africa, however, a considerable proportion of women are not living with their biological children; and this is associated with high levels of child fosterage. In this section I explore the cultural, political and socio-economic factors which have influenced the residential patterns of women and children. I also consider how these residential patterns create a unique environment for a study of the effects of children on women's labour force participation.

South African households are characterised by complex living arrangements. The co-residence of extended families in African households, especially in rural areas, is not unique to South Africa alone. This cultural practice also has historical roots in other developing countries, particularly in Africa. However what is unique to the South African context is that political arrangements have significantly impacted on the form and structure of households (Posel and van der Stoep 2008; Amoateng *et al* 2007; Zulu and Sibanda 2005).

The apartheid system considerably influenced residential patterns as it placed restrictions on where and with whom Africans could live (Posel and van der Stoep 2008). The permanent settlement of Africans and their families in urban areas was prohibited under apartheid laws (Posel and Casale 2006; Posel 2002). This resulted in 'circulatory' or 'temporary' patterns of labour migration to places of employment, particularly from rural to urban areas. Because migrants could not settle permanently in urban areas and could not move with their families, they would retain a base in their household of origin to which they would return at some future date. In the case of female migration, women could not move freely with their children and were forced to leave their children in the care of parental surrogates in the household of origin.[6] For example, female domestic workers, who comprised a large component of African female migrants, were prohibited from living with their children in White residential areas under the Group Areas Act (Zulu and Sibanda 2005:237).

Although controls on urbanisation no longer exist post-apartheid, patterns of temporary labour migration still persist. In fact there has been a reported increase in female labour migration in recent years (Posel and Casale 2006, 2003; Collinson *et al* 2003; Collinson and Wittenberg 2001). Using nationally representative household surveys, Posel and Casale (2006; 2003) find that between 1993 and 2002 an increasing proportion of rural African households reported at least one labour migrant where this increase has been driven specifically by female migration. Collinson *et al* (2003) and Collinson and Wittenberg (2001), in a smaller study of labour force dynamics in the rural Agincourt district of the Northern Province, find a marked upturn in female labour migration from 1997 to 2001.

With the removal of restrictions on the mobility of migrants with their families, it is expected that women would migrate with their children. On the contrary, evidence suggests that female migration is still associated with the fragmentation of mother-child residential patterns (Anderson and Phillips 2006; Collinson *et al* 2003). Anderson and

[6] It must be noted, however, that female migration was more constrained than male migration in apartheid South Africa, but among women who were migrants, movement with children was prohibited.

Phillips (2006:15) report that a significant proportion of African children whose mothers were temporary labour migrants[7] were fostered out from 1995 to 1998. There are various possible reasons for this. Accommodation for children may not be readily available at places of employment and living conditions in urban centres may not be considered suitable environments for raising children (Posel and Casale 2006; James 2001). Furthermore, childcare may be unavailable or too expensive in urban areas, compared to rural areas where the co-residence of extended families provides informal 'affordable' childcare.[8]

The evidence above suggests that labour migration continues to play an increasingly important part of female labour force participation in South Africa. Furthermore, mothers continue to migrate without their children, leaving them behind in the household of origin where the co-residence of extended families increases the availability of parental surrogates (Posel and van der Stoep 2008). Where mothers are not co-resident with their children and thus do not face direct childcare constraints, they may have more opportunity to pursue labour market opportunities than mothers who are co-resident with their children (*ibid*). The remaining sample of co-resident mothers are therefore expected to be a non-random sample of all women who are less likely to participate in the labour force.

With the fragmentation of family structure under apartheid laws, it is no surprise that child fosterage has historically been a common occurrence in African households (Amoateng *et al* 2007; Bray and Brandt 2007; Anderson and Phillips 2006). Customary practices and traditions have also contributed to child fosterage in South Africa. Studies

[7] Anderson and Phillips (2006:15) identify women who are likely to be temporary labour migrants as mothers who were neither household heads nor a spouse to the household head in the destination household.

[8] Research by Bray and Brandt (2007) also suggests that residential mobility is perceived as being detrimental to child development among African caregivers. They note that "caregivers who participate in high rates of residential mobility emphasise the importance of stability in the home setting for good-quality child care" (Bray and Brandt 2007:6). In practice, however, oscillatory patterns of child residency are very common in South Africa (Van der Waal 1996).

have shown that African parents have circulated their children among relatives as a way of strengthening families ties or for better schooling (or skills training) opportunities that cannot be obtained in their area of residence (Bray and Brandt 2007; Collinson *et al* 2003; McDaniel and Zulu 1996; Van der Waal 1996). In contemporary South African patterns of child fosterage have continued for various reasons. Along with persistent patterns of female migration, Van der Vaal (1996) finds that another reason for African parents fostering out their children is an inability to support these children. Increasing AIDS-related deaths or ill-health among mothers has also resulted in an increase in the percent of fostered children in recent years (Anderson and Phillips 2006). Noumbissi *et al* (2005), for example, find that when a mother dies it increases the odds that children will be sent to live with relatives, especially grandmothers, in an extended family structure. There is also evidence to suggest that in households with HIV/AIDS infected individuals, mothers may even leave their households due to overwhelming care burdens, abandoning their children in the process (Bray and Brandt 2007; Denis and Ntsimane 2006).

High levels of child fosterage and the co-residency of extended families in South Africa may have implications for a study of the effect of children on women's labour force participation. On the one hand this affordable childcare may lower the price of a mother's time in non-market work. On the other hand non-mothers, who act as parental surrogates, may face a higher price of time in non-market work as childcare demands compete with their time in market work. Therefore caring for children who are not their own may also reduce the probability of labour force participation among non-mothers (Posel and van der Stoep 2008). I explore these implications in Chapters Four and Five.

Conclusion

In this chapter both theoretical and empirical literature on the relationship between childbearing and labour force participation have been reviewed. I first considered economic theory of labour supply with the view to understanding how childbearing/fertility behaviour influences labour force participation decisions. Both the

traditional neoclassical model of labour supply and Becker's theory of the allocation of time, suggest that childbearing lowers the probability that women participate in the labour force. The mechanism by which this occurs is that childbearing raises the relative price of a mother's time in non-market work above her potential wage rate, so that she is less likely to supply any amount of time to the labour market. The mother's price of time in non-market work will vary by the age and number of children she has. Young children, specifically, will raise the value of her price of time because they require intensive supervision.

The theory, however, suggests that a two-way causal relationship exists between childbearing/fertility and labour force participation. Labour force participation is itself a determinant of women's fertility behaviour. Furthermore, mothers may be different from non-mothers in terms of their unobserved characteristics, such as ambition or motivation, which may be negatively correlated with childbearing but positively correlated with the decision to participate in the labour force. Therefore, the fundamental problem in estimating this relationship is the endogeneity of childbearing in labour force estimations. Instrumental variable estimation has been the most commonly used approach to address endogeneity issues. I provided a detailed review of these international studies and the unique strategies used to identify the exogenous effects of childbearing on women's labour force participation.

Even after controlling for endogeneity in motherhood, studies from developed countries consistently find a negative relationship between childbearing and labour force participation; and this relationship is strongest among women with very young children. However evidence of a negative relationship is less robust in developing countries. A negative, positive or no relationship is found. Differences in the nature of work and the social organisation of childcare largely account for this ambiguous relationship in developing countries.

In the South African context, in particular, the social organisation of childcare is likely to have implications for a study of the effects of childbearing on labour force

participation. The co-residence of extended families among African households is likely to increase the availability of parental surrogates resulting in a network of informal childcare. On the one hand, the presence of affordable childcare may help mothers combine work and childrearing commitments. On the other hand, as childcare responsibilities are borne by non-mothers it may reduce the probability that they participate in the labour force. This is likely to be a common occurrence given high levels of fosterage in South Africa.

An important part of labour force participation in South Africa is migration to places of employment. Despite the removal of controls on African urbanisation in post-apartheid South Africa, female migration has been increasing in recent years; and there is evidence to suggest that women continue to migrate without their children. A significant sample of mothers is therefore likely not to be resident with their children. These not co-resident mothers may face more opportunity to participate in the labour force than co-resident mothers because their time is unconstrained by day-to-day childrearing activities. In the international literature on the relationship between motherhood and labour force participation, a co-residency requirement is typically placed on definitions of motherhood. But in the context of South Africa, the remaining sample of co-resident mothers is likely to be a non-random sample of all mothers who are less likely to be labour force participants. A co-residency requirement is therefore likely to bias the estimated relationship between motherhood and labour force participation in South Africa.

With the exception of Mlantsheni and Leibbrandt (2001), studies have not explored if there is a motherhood/childbearing effect on women's labour force participation in South Africa. A possible explanation for this is that studies on female labour force participation in South Africa have not been able to match women to their children with datasets that have been analysed. The majority of household surveys that collect information on labour force activity do not simultaneously collect detailed birth history information. In the following chapter, I analyse which data are available identify mothers and the quality of these data.

Chapter Two:

Identifying mothers in South Africa – data and methods

Introduction

The previous chapter highlighted that studies on female labour force participation in South Africa have paid scant attention the effects of childbearing. A reason for this is that studies have not been able to match women to their children using household surveys. Although a vast literature exists on identifying fertility rates (see for example Palamuleni *et al* 2007; Moultrie and Timaues 2002, 2001; Mencarini 1999; Udjo 1998; Mostert and Hofmeyer 1988), little or no literature has identified or measured mothers in South Africa. A first step to estimating the relationship between motherhood and labour force participation in South Africa is to identify among women who are mothers. The objective of this chapter is to investigate available methods of matching women to their children using household surveys in South Africa. The chapter also analyses the quality of the available data for identifying women with children.

Motherhood is most commonly referred to as a biological phenomenon. In the economics literature, in particular, mothers are typically identified as women with biological or own children and who are co-resident with these children (Boushey 2008; Vere 2008; Cruces and Galiani 2007; Angrist and Evans 1998). But in the South African context of child fosterage, extended household living arrangements and labour migration, mother and child relationships may extend beyond biological and spatial boundaries (Bray and Brandt 2007). In this context I present four different methods available for identifying mothers using household survey data in South Africa. I contrast and compare these different methods, highlighting how each presumes different definitions of motherhood. I also point out their potential benefits and limitations in identifying mothers.

These four identification methods have particular relevance for the rest of this thesis. First, they provide a basis for measuring mothers, results of which are presented in Chapter Three. Second, the various methods of identifying mothers have significant implications for the estimation of the relationship between motherhood and labour force participation (see Chapters Five and Six).

Identifying mothers requires datasets which can be used to match woman to their children. But this is only possible with a select number of South African household surveys. A large section of this chapter considers which cross-sectional as well as panel datasets are available to identify mothers in South Africa. I analyse the quality of these data where possible and highlight that birth data are compromised by high levels of non-response on children ever born, measurement error and sample attrition. I also consider a method developed by El-Badry (1961) to adjust or correct for incomplete or missing birth status entries. El-Badry argues that a large proportion of women with missing birth status information are in fact childless women. I apply the El-Badry adjustment to the Population Census 1996 data to estimate the extent to which missing birth responses are overstated and childless women understated.

In the next section I discuss definitions of motherhood in the developing country context. In section 2.2 I consider data available to match women to their children. Section 2.3 discusses the four methods of identifying mothers using household surveys. Sections 2.4 and 2.5 analyse the quality of birth data in cross-sectional and panel surveys respectively. Finally, I discuss and apply the El-Badry adjustment to the Census 1996 data in section 2.6.

2.1 Defining motherhood in developing countries

Definitions of motherhood vary across developed and developing countries. In developed countries, characterised by nuclear family structures, motherhood typically refers to women with biological ties to their children. In the economics literature, in particular, mothers are usually identified as women who have had a live birth and at

least one of these children is aged 18 years or younger (Boushey 2008; Cruces and Galiani 2007; Angrist and Evans 1998; Bronars and Grogger 1994). Furthermore these studies impose a *co-residency requirement* on motherhood (Boushey 2008; Vere 2008; Cruces and Galiani 2007; Angrist and Evans 1998; Bronars and Grogger 1994).

In the developing country context, however, it may be more appropriate to expand the definition of motherhood to include mothers who do not co-reside with their children as well as women who care for children who are not biologically their own. A woman may become a mother through processes that are not biological; for example through adoption, child fosterage or even assuming or being assigned the role of a parental surrogate (Russell 2004; Sudarkasa 2004). This is particularly the case in Africa and South Africa specifically, where childcare responsibilities extend beyond the biological mother to the wider community (Bray and Brandt 2007; Russell 2004; Sudarkasa 2004).

The previous chapter highlighted that a large proportion of South African households are characterised by extended family structures. Furthermore, cultural traditions of 'circulating' children among relatives coupled with increased female labour migration and the HIV/AIDS epidemic have resulted in high levels of child fosterage among Africans in South Africa (Amoateng *et al* 2007; Bray and Brandt 2007; Van der Waal 1996). In this case, childcare is often a communal activity rather than isolated to the biological mother of the child (Bray and Brandt 2007; Mkhize 2006; Sudarkasa 2004). Individuals may assume childcare responsibilities for other children in the household who are not their own. The result is that the term "mother" is often blurred across biological boundaries and may be used to refer to a range of other individuals such as grandmothers, aunts or even older siblings (Ramphele and Richter 2006; Mkhize 2006; Anderson and Phillips 2005; Sudarkasa 2004; Magwaza 2003; Blacker 1977). Mkhize (2006) notes that the extension of terms such as mother to other individuals in the household, "pledges these people to parental responsibilities" (*ibid*:187). Studies that identify the effects of children on women's labour force participation in South Africa, should attempt to identify women who live with children who are not their own and have assumed the responsibilities and behaviours typical of a mother.

The effect of female labour migration in separating mothers from their children adds a particularly interesting dimension to the analysis of motherhood in South Africa. The responsibilities in caring for children differ across mothers who do and do not co-reside with their children. In the case of co-residence, motherhood responsibilities involve both caring for the physical and emotional needs of their children as well as their financial needs. In the not co-residence case, there are no direct childcare responsibilities but motherhood may involve a financial burden in caring for a child's well-being. In fact, providing for children may be a key motivation for mothers moving away from their children to work or to search for work. Excluding not co-resident mothers from measures of mothers would ignore the relationship between motherhood and labour force participation for these women. Therefore measures of mothers in South Africa should identify both sets of mothers but also distinguish them on the basis of co-residency with children.

2.2 Data: matching women to their children

There are a number of sources of data in South Africa on individuals and the households in which they live; but it is not always possible to match women to their children in these household surveys. In particular, if motherhood is understood as a biological phenomenon then only a few select household surveys that collect information on fertility can be used to identify women with and without children. The first of these is the Project for Statistics on Living Standards and Development (PSLSD) conducted in 1993. This is the first nationally representative survey in South Africa which was coordinated and managed by the South African Labour Development Research Unit (SALDRU). The PSLSD interviewed approximately 9,000 households and contains birth histories of women aged 15 to 49. Since 1993, a select number of October Household Surveys, General Household Surveys and the two Population Censuses, conducted and managed by Statistics South Africa, have also collected birth information. The October Household Surveys (OHS) interview approximately 30,000 households in South Africa. Specifically the 1995, 1997 and 1998 OHS's included a birth module which questioned

women aged 12 to 55 on their birth histories. The OHS 1999 included a module on births but it only questioned women on their birth histories in the past 12 months (Casale 2003). The October Household Surveys were replaced by the Labour Force Survey (LFS) which has been conducted biannually by Statistics South Africa since February 2000. To date, none of LFS's has questioned respondents on their birth histories. The 1996 and 2001 Population Censuses both included a question on the number of children ever born to a woman in the household between the ages of 12 and 50.

The most current nationally representative household survey, which contains detailed birth histories, is the General Household Survey (GHS) 2002. The birth module in the GHS 2002 questioned all women aged 12 to 50 on children ever born, the order of births and the sex and age of each child born. The GHS 2002 also surveyed approximately 30,000 households.

There are other household surveys which do not include any birth history information but contain either a question on the relationship of individual household members to the household head or a maternal relationship question. These relationship questions can be used to link women to their children. In this case a necessary requirement to identify a mother is that children are *resident in the same household* as the mother. With the exception of the LFS's, all the aforementioned surveys contain relationship questions to match mothers to their children. In addition the OHS's 1997 to 1999 contain relationship questions as well as the GHS's 2003 to 2006.

The surveys mentioned above are cross-sectional surveys where information is collected on individuals or households at a single point in time. Panel data sets, however, collect information on the *same* individuals over *different* points in time. There are a few panel datasets available in South Africa which contain birth modules used to identify mothers: The KwaZulu-Natal Income Dynamics Study (KIDS); the Cape Area Panel Study (CAPS); the Africa Centre Demographic Information System and population-based HIV Survey (ACDIS) and Agincourt Health and Demographic Surveillance System (AHDSS). The KwaZulu-Natal Income Dynamics Study tracked a sample of African

and Indian households in KwaZulu-Natal in 1998 and 2004 that were originally surveyed in the 1993 Project on Living Standards and Development (May *et al* 1999). The Cape Area Panel Study tracked a sample of youths and young adults only in the Metropolitan Cape Town area over four waves from 2002 to 2006 (Lam *et al* 2006). The ACDIS panel began in 2000 and tracks a sample of predominately Zulu-speaking individuals in the rural Umkanyakude district of KwaZulu-Natal (Tanser *et al* 2007). AHDSS tracks a sample of households in Agincourt, in the rural Bushbuckridge district near the Mozambique border in the Limpopo Province (University of Witwatersrand 2008). Unfortunately the ACDIS and AHDSS datasets are not in the public domain.

None of the aforementioned panel surveys are nationally representative. The forthcoming National Income Dynamics Study, conducted by SALDRU, will be the first nationally representative panel study in South African to identify women's birth status (National Income Dynamics Study 2008). The first wave of this panel study of 8,000 households has been conducted recently in 2008. In subsequent years, this dataset will provide very useful information to identify and track the motherhood status of women over time.

In many household surveys in South Africa the quality of birth information is compromised by high levels of missing data or non-response which does not bode well for obtaining accurate counts of mothers. After suggesting different methods for identifying mothers in South Africa, I analyse the quality of birth information in the following cross-sectional surveys: the PSLSD 1993; the OHS 1995 and 1997; the Census 1996 and 2001 and the GHS 2002. I also analyse problems of missing birth data, measurement error and panel attrition in KIDS using waves one and two (1993 and 1998).

2.3 Four methods for identifying mothers using nationally representative household surveys

In South Africa motherhood could be defined as merely a biological occurrence or as a cultural practice that extends beyond biological ties to other women. A variety of

methods to identify mothers are therefore possible. I suggest two sets of ways to identify mothers using the information available in South African household surveys. The first set (referred to as methods one and two) imposes a biological restriction on motherhood and requires fertility data to identify mothers. The second set (referred to as methods three and four) removes a biological restriction on motherhood and is based on household relationship questions. Both sets do not restrict the definition of motherhood on the age of children.

Considering the first set of identification methods I propose two ways to identify mothers using fertility data. Method one identifies a woman as a mother if she has had at least one live birth, regardless of whether the child is still alive. Use of this method requires a survey containing a birth module. Where birth modules are available, they are typically asked of resident women in the household aged 12 to 50 (or 55 in some cases) and contain a standard question on whether a woman has had a live birth. Birth modules also typically ask whether the child born to the mother is still alive. Using this question, a second method of identifying mothers further restricts mothers to women who have biological children who are still alive.

A major benefit of using birth data to identify mothers is that it identifies biological mothers regardless of whether they live with their children. Methods one and two are able to capture the motherhood status of women who live away from their children for reasons such as labour migration. But using these methods to measure mothers may be compromised by high levels of non-response on children ever born, especially among adolescents. I explore this in more detail in section 2.4 and 2.5. Another key aspect of these methods is that they restrict motherhood to women whose children are biologically their own. Women with adopted or fostered children who act as parental surrogates to other children in their household should not be captured here as mothers.

Considering the second set of identification methods, commonly available questions on household relationships can be used to match women to their children. Many surveys include a maternal relationship question which asks an individual whether his/her

mother is alive and the person code of his/her mother *if she resides in the same household.* If at least one individual provides the person code of a woman in the household, then the woman in question is identified as a mother.

In the PSLSD 1993, the OHS 1995, the Census 1996 and the GHS 2002 and 2003, the wording of this maternal relationship question does not restrict a respondent's reply to his or her biological mother (see Appendix 2A for questions). Blacker (1977:108) notes that in African households "the words 'father' and 'mother' are often used loosely to denote not only a person's biological parents, but also foster parents, or older relatives acting, perhaps temporarily, in *loco parentis*, or simply as terms of respect for members of an older generation". For this reason method three may capture adopted or foster mothers or even women who are considered to play the role of a maternal caregiver in the household. But in the later OHS surveys (1997 to 1999), the Census 2001 and the GHS 2004 to 2006 the maternal relationship question was modified by restricting an individual's response to his/her biological mother. Applying method three to these surveys may not capture non-biological mothers.

A benefit of method three is that it captures motherhood status for women over childbearing age. This may be particularly useful in South Africa where the elderly are increasingly bearing the burden of childcare in the context of the HIV/AIDS epidemic. Another benefit is that it is based on questions less subject to non-response compared to the first set of identification methods. There are two reasons for this. First, the question on maternal relationship is not asked in the sensitive context of fertility which diminishes problems of non-response associated with birth status questions. Second, it is not directed to the woman in question but rather to her children who may be less reluctant to answer questions on maternity.

However a major limitation of method three is that it restricts mothers to women who live in the same household as at least one of their children. Identifying mothers in this manner will capture women who provide for the physical or emotional care of children with whom they reside. The co-residency requirement, however, excludes from the

measure of mothers women who live away from their children, perhaps to earn a living to support their families. In the context of high levels of female labour migration in South Africa, this method may fail to identify a sizeable sample of mothers who are not co-resident with their children.

The fourth method of measuring mothers uses information collected in household surveys on kin relationships to the head (or acting head) of the household. I follow the methodology of Posel and Devey (2006) who use information on kin relationships as one way in which to identify *fathers* in South Africa. Adapting Posel and Devey's application of questions on household kin-relationships, a woman could be identified as a mother if the woman is:

1) a female head or wife/partner to the male head *and* an individual in the household reports being a son/daughter of the head;
2) a parent to the head;
3) a grandparent to the head;
4) a female head or a wife/partner to the male head *and* an individual reports being a grandchild of the head;
5) an adult daughter to the head *and* an individual reports being a grandchild to the head.

At first glance, the fourth identification method resembles the third method. These similarities are outlined in summary Table 2.1 which highlights and compares the key aspects each method. However method four is expected to yield the lowest measures of mothers compared to the previous methods. Posel and Devey (2006:40-41) report much lower proportions of men who are *fathers* using this method compared to the results of other methods that link fathers to their biological children. The first reason for underreporting mothers using method four is that the available relationship options inadequately capture the motherhood status of women living in extended households who are not spouses or parents of the head of the household. For example, it would not be possible to identify a mother who reports being a sister or cousin to the household head. It follows that mothers living in laterally extended households would be undercounted.

Method four is also subject to misclassification. For example, there is no certainty that a grandchild in the household is the biological child of an adult daughter (the fifth criteria above) (Posel and Devey 2006:41). This would result in overestimating mothers. Another reason for misclassification of women with and without children is that in many African households, kinship terms used in English which refer to close relatives may be applied to distant relatives in the extended family (Sudarkasa 2004). Relationship terms may be interpreted differently by the respondent so the individual's reported relationship to the household head is different from his or her *de facto* relationship. Comparing counts of mothers over time using these relationship codes is also problematic since the relationship options provided vary across the different surveys. This is especially the case with regards to the treatment of stepchildren, adopted children, in-laws, cousins, aunts and uncles. Appendix 2B contains the relationship question in the PSLSD 1993, the OHS 1995 and the GHS 2002 and the associated relationship responses. It shows variation in relationship responses across the three surveys that complicates the comparison of measures of mothers over time.

It is also possible to define motherhood using combinations of the four methods. The most inclusive definition of motherhood that is possible with household survey data combines method one, three and four. A woman is a mother if she:

1) has either had a live birth *or*
2) is reported as living together with at least one child *or*
3) is defined as a mother through household kin-relationships.

Although this provides an upper limit estimate of motherhood using available data in South Africa, it may still undercount non-biological mothers who do not live at home with their children. For example, it will not capture women who have adopted children but these children are attending school in another area and therefore are not living with their adopted parents.

Table 2. 1: A summary of the four methods to identify mothers using household surveys

	Method one: Woman has had a at least one live birth	**Method two: Woman has had a live birth *and* at least one child is still alive**	**Method three: Woman lives in the same household as an individual who reports her to be his/her mother**	**Method four: Using information on kin-relationships to the household head**
Derived using information from	Birth modules	Birth modules	Relationship question on who a person's mother is in the household	Question on kin-relationship to the household head
Biological / non-biological motherhood	Biological	Biological	Biological and non-biological	Biological and non-biological
Motherhood restricted to women whose children are alive	No	Yes	Yes	Yes
Counts include mothers who do not live with their children	Yes	Yes	No	No
Age limit	Childbearing age	Childbearing age	All ages	All ages
Levels of non-response	High (especially among young age cohorts)	High (especially among young age cohorts)	Very low	Very low
Comparability across household surveys	Good	Good	Fair	Poor
Nationally representative household surveys to which this method applies	PSLSD 1993 OHS 1995 OHS 1997 Census 1996 OHS 1998 Census 2001 GHS 2002 *Not* LFS's	PSLSD 1993 OHS 1995 OHS 1997 Census 1996 OHS 1998 Census 2001 GHS 2002 *Not* LFS's	PSLSD 1993 OHS 1995 GHS 2002 GHS 2003 *Not* OHS 1997 - 1999 *Not* Census 2001 *Not* GHS 2004-6 *Not* LFS's	PSLSD 1993 OHS 1995 Census 1996 OHS 1997 - 1999 Census 2001 GHS 2002 - 2006 *Not* LFS's

2.4 Methods one and two: missing data in cross-sectional surveys

In this section I show that in cross-sectional household survey datasets which contain birth modules, the quality of birth information is compromised by high levels of missing data. I can determine levels of missing birth data in the PSLSD 1993, the OHS 1995 and the ten percent sample release of the Census 1996; however, it is not possible to distinguish childless women from women with missing birth data in other household survey datasets, namely the 1997 OHS, the ten percent sample release of the Census 2001 and the GHS 2002 for reasons that are explained later in this section. Therefore I limit the following analysis of missing birth data to the PSLSD 1993, the OHS 1995 and the Census 1996.

Table 2.2 identifies the proportion of women aged 20 to 49 with missing data on birth status. The level of missing data on children ever born is lowest in the PSLSD 1993 at about five percent compared to 11 percent in the OHS 1995 and about ten percent in the Census 1996. There are also clear racial and age dimensions to these results. Levels of missing birth data are highest among White women in the PSLSD 1993, but in the OHS 1995 and the Census 1996 it is highest among African women. In Table 2.3 and Figure 2.1 I identify proportions/percentages of women with missing birth data in each age cohort, *including* age cohort 15 to 19. In all three surveys, missing data are most pronounced in the youngest age cohorts, particularly ages 15 to 19. In the 1995 OHS, for example over forty percent of women aged 15 to 19 have missing birth status information compared to 24 percent of women aged 20 to 24.

In the PSLSD 1993, the OHS 1995 and the Census 1996 the incidence of missing birth data drops significantly among the 20 to 24 age cohort compared to the 15 to 19 age cohort and continues its decline to age 49. At ages 30 to 49 missing birth data fall to low levels even in surveys with the highest overall levels of missing birth data. This is illustrated clearly in Figure 2.1. The gaps between the three line plots show that differences in birth data quality across surveys are most evident at younger age cohorts but are eliminated at older cohorts. Therefore an important criterion in comparing the quality of birth data across South African surveys is the level of non-response among women aged 15 to 19.

Table 2. 2: Proportion and counts of women aged 20 to 49 with missing birth status

	PSLSD 1993	OHS 1995	Census 1996
All	0.045 (0.003) 370,566	0.115 (0.004) 991,544	0.102 (0.000) 876,342
African	0.032 (0.003) 193,773	0.136 (0.004) 887,730	0.110 (0.000) 708,357
Coloured	0.047 (0.008) 35,026	0.048 (0.003) 40,733	0.067 (0.001) 55,253
Indian	0.028 (0.009) 7,199	0.078 (0.011) 19,901	0.096 (0.002) 24,037
White	0.112 (0.014) 134,568	0.042 (0.002) 43,180	0.079 (0.001) 77,355

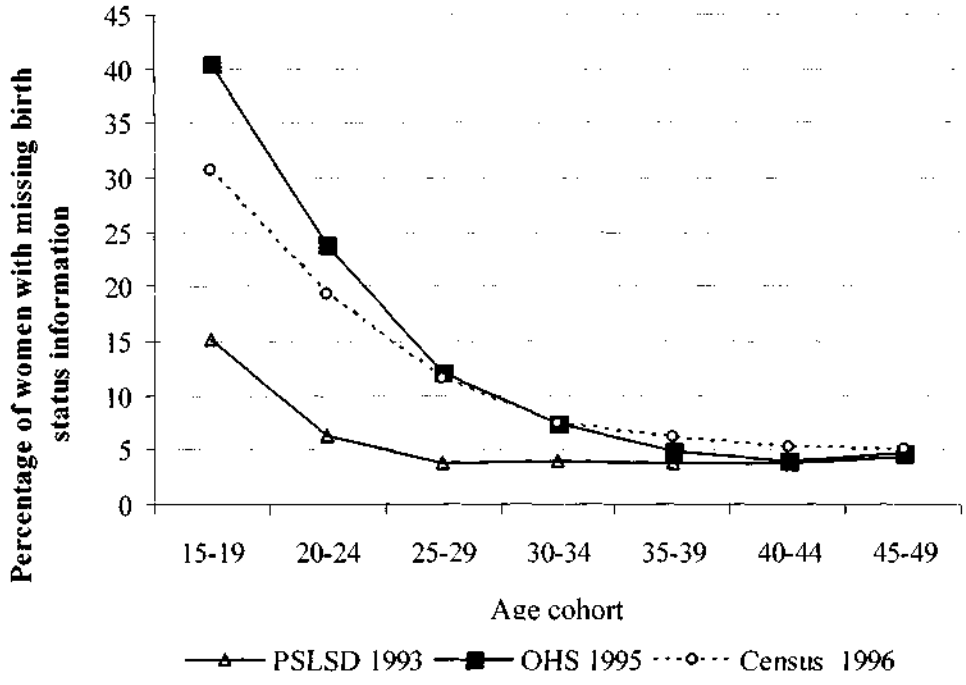
Notes: Standard errors are in parentheses. The bottom figures in each cell are weighted counts. Estimates account for weighting and clustering in sample survey designs. Missing birth status is calculated for a sample of women aged 20 to 49 who are resident household members. Missing birth status is indeterminate in the OHS 1997, the GHS 2002 and the Census 2001 because it is impossible to distinguish women who have had zero births from women with missing birth status.

Table 2. 3: Proportion and counts of women with missing birth status by age cohort

Age	PSLSD 1993	OHS 1995	Census 1996
15 – 19	0.151 (0.010) 333,727	0.403 (0.008) 852,894	0.308 (0.001) 623,759
20 – 24	0.063 (0.006) 127,249	0.243 (0.008) 511,010	0.193 (0.001) 389,235
25 – 29	0.038 (0.005) 61,901	0.122 (0.005) 213,753	0.116 (0.001) 204,645
30 – 34	0.039 (0.005) 55,624	0.076 (0.005) 121,965	0.075 (0.001) 121,461
35 – 39	0.038 (0.007) 46,541	0.049 (0.005) 64,738	0.061 (0.001) 83,779
40 – 44	0.038 (0.007) 41,536	0.039 (0.005) 41,622	0.053 (0.001) 58,586
45 – 49	0.044 (0.007) 37,715	0.047 (0.004) 38,454	0.051 (0.001) 44,033

Notes: Standard errors are in parentheses. The bottom figures in each cell are weighted counts. Estimates account for weighting and clustering in sample survey designs. Missing birth status is calculated for women who are resident household members. Missing birth status is indeterminate in the OHS 1997, the GHS 2002 and the Census 2001 because it is impossible to distinguish women who have had zero births from women with missing birth status.

Figure 2. 1: Percent of all women aged 20 to 49 with missing birth status information



Source: PSLSD 1993, 1995 OHS and Census 1996.
Notes: Standard errors are in parentheses. Results account for weighting and clustering in sample survey designs. Missing birth status is calculated for women who are resident household members. Missing birth status is indeterminate in the OHS 1997, the GHS 2002 and the Census 2001 because it is impossible to distinguish women who have had zero births from women with missing birth status.

One reason for high levels of non-response among the younger age cohorts may be the reluctance among girls to reveal sensitive information on sexual partnerships to other family members, especially parents (Preston-Whyte 2003; Makiwane 1996). Preston-Whyte (2003:90) notes that in the context of families there may be " 'silences' over sex and reproduction between generations and genders". Where sex is shrouded in secrecy, knowledge of adolescent pregnancy may be concealed from other household members. Another reason which is commonly documented, especially in developing country surveys, is that enumerators fail to record a zero entry for the number of children ever born when a woman is childless (El-Badry 1961). This has been noted as a problem in South African birth data (Moultrie and Dorrington 2004; Moultrie and Timaues 2003). The result is that missing birth entries are recorded for women who are in fact childless (i.e. non-mothers). This reporting error is termed "zero-error" by El-Badry (1961). In younger age cohorts, there are greater numbers of childless women compared to older age cohorts, increasing the

incidence of zero-error reporting in younger age cohorts i.e. higher proportions of women with missing birth information who should be captured as childless women (El-Badry 1961).

In order to avoid the confounding effects of high levels of missing data among younger age cohorts on research conclusions, this thesis limits its analysis to women aged 20 to 49. If reliable measures of motherhood are to be obtained for a younger sample, then in the future design and execution of surveys more careful consideration needs to be given to reducing fertility non-response. South African national household surveys should explore new ways to question adolescents or young adults on their birth histories separately from older adults. For example, in the Cape Area Panel Study (CAPS) interviewers were instructed that the fertility module be administered in private with each young adult or that the questionnaire be completed directly by the respondent if he/she was "uncomfortable answering the questions out loud or if other people were present" (Lam *et al* 2006: 28). Furthermore, more effort should be taken to train enumerators in administering birth module questions in order to prevent zero-errors in reporting.

It is not possible to distinguish childless women from women with missing birth entries in the 1997 OHS, the Census 2001 and the GHS 2002. There are two reasons for this. First, questionnaire design and enumerator prompts often do not allow missing responses and childless responses (zero birth entries) to be distinguished as two separate categories. Second, imputation of missing data prior to public release of datasets inhibits analysis of missing data which have already been replaced by imputed values.

In the OHS 1997 and the GHS 2002 survey design rather than imputation is the primary reason for not being able to distinguish women with missing birth status from childless women. These surveys are designed in such a way that positive data entries are only made for women who had ever given birth and no zero entries can be made for childless women. Therefore childless women who would have reported zero births are indistinguishable from women who would not respond to the birth question as they are all coded as having missing birth information. In the Census 2001, however, the questionnaire was designed to

distinguish these two groups of women because both positive, zero and missing birth entries could be made.

However, in the ten percent sample release of the Census 2001 the inability to distinguish missing responses from childless responses is primarily the result of imputations for missing data by Statistics South Africa prior to the public release of the data. Statistics South Africa reports an imputation rate of 17.4 percent for the unknown total live birth status of females between the ages of 12 and 50 (Statistics South Africa 2002). (For detailed analysis of data quality and the use of imputation for birth data in the Census 2001 see Moultrie and Dorrington (2004).)

The inability to determine levels of missing data in household surveys aggravates problems of estimating motherhood. It becomes impossible to determine to what extent mothers (or childless women) may be underestimated (or overestimated) without knowing the level of missing birth data or non-response. It also becomes difficult to make conclusions about changes in the extent of motherhood over time when comparing survey data.

2.5 Methods one and two: missing data and measurement error in panel data

There are many advantages of using panel data over cross-sectional data in demographic studies. Panel surveys collect information on the same individuals or households at multiple points in time. This enables the researcher to observe dynamic processes such as entries into motherhood and birth spacing of children. Furthermore panel data is useful for controlling for endogeneity in estimations of the effect of motherhood on labour force outcomes. A key source of endogeneity bias is that unobserved characteristics which are correlated with both a woman's childbearing behaviour and her labour force activity are included in the error term in an estimation equation. Using panel data, unobserved characteristics that remain constant over time can be differenced out of an equation because multiple data points are observed on the same individual. Thus endogeneity sources can be eliminated allowing the researcher to estimate an exogenous effect of motherhood on the

outcome variable.[9] (For more detailed information on the advantages of panel data see Hsiao 2003.) Despite the advantages of using panel data over cross-sectional data, the quality of birth data may be further compromised in panel datasets. Samples suffer from attrition where individuals fall out of the panel over time either through death, migration or refusal to participate in the repeated survey (Alderman *et al* 2001:83). Measurement error in birth data also becomes more pronounced when reconciling conflicting responses on birth status at different points in time. For example, a woman may report having ever had a birth in one year but in the next year the same woman reports a zero birth entry.

To illustrate problems of birth data quality in panel surveys, I use the KwaZulu-Natal Income Dynamics Study (KIDS), waves one (1993) and two (1998). KIDS collected quite comprehensive birth histories on resident women aged 15 to 49. I have focused my analysis on KIDS for the following reasons. First, unlike the ACDIS and the AHDSS, KIDS is in the public domain. CAPS is also in the public domain but it only collects detailed birth information on adolescents and young adults and therefore is not useful for identifying all adult mothers. The KIDS dataset also contains a comprehensive set of data on socio-economic variables and therefore would be useful in a study of motherhood and its economic outcomes. But the analysis to follow suggests that poor data quality in KIDS may lead to biased results in such a study.

The second wave of KIDS resurveyed 1,393 African and Indian households in KwaZulu-Natal in March-June 1998 that were originally interviewed in the 1993 PSLSD. KIDS collected information on resident and non-resident household members aged 15 to 49. Restricting the sample to African women aged 15 to 44 in 1993, I identify 1,767 resident women.

[9] See for example Anderson *et al* 2002; Joshi 2002; Budig and England 2001; Waldfogel 1997 who use panel data to estimate the exogenous effects of motherhood on a woman's earnings potential.

Table 2. 4: Tracking women across the first two waves of the KIDS panel

1993 Resident African woman, aged 15-44 N = 1,767								
% located in 1998 ALIVE 85.12 (1,504)			% located in 1998 DEAD 2.09 (37)			% not located in 1998 12.79 (226)		
Birth status in 1993, %:			**Birth status in 1993, %:**			**Birth status in 1993, %:**		
No births	**At least one birth**	**Missing birth status**	**No births**	**At least one birth**	**Missing birth status**	**No births**	**At least one birth**	**Missing birth status**
34.71 (522) ↓	61.12 (921) ↓	4.06 (61) ↓	21.60 (8)	75.68 (28)	2.70 (1)	33.19 (75)	61.95 (140)	4.87 (11)
Percent of women remaining resident / *non-resident* in the household								
64.94 (339) *33.15* *(183)*	69.92 (644) *30.08* *(277)*	67.21 (41) *32.79* *(20)*						

Source: KwaZulu Natal Income Dynamics Study (KIDS) 1993 and 1998.
Notes: In the first row of figures, percentages sum up to 100 percent in each subdivision. In the second row of figures, percentages are calculated as a total of the observation number in the directly preceding category above. Observation numbers are in parentheses.

Unfortunately, KIDS birth data are subject to two types of panel attrition. First, women can fall out of the panel if they live in households which could not be located in 1998 or if they had died since 1993. This represented 15 percent ([226+37]/1,767) of women in the original 1993 sample. The remaining 85 percent (1,504/1,767) were resurveyed in 1998.

Maitra and Ray (2003:1041) contend that a resurvey rate of 85 percent is quite satisfactory for a developing country panel data set. But with respect to tracking women's fertility, a second type of attrition occurs in KIDS that exacerbates sample attrition. The survey does not collect birth information on non-resident household members. Resident women in 1993 who become non-resident in 1998 are no longer counted in the sample of women. As a result birth status could not be observed for 32 percent (480/1,504) of women who were resident household members in 1993, and were still alive in 1998, but had moved away from the household. Given the two types of attrition it is possible to track the birth status of only 1,024 (58 percent) of the original 1,767 women. Panel attrition in 1998 is therefore over forty percent of the original sample in 1993.

An overall attrition rate of over forty percent for a sample of women could be considered as being high relative to attrition rates in other developing country panel surveys. Alderman *et al* (2001:81), in a study of sample attrition in panel data sets in other developing countries[10], document that overall attrition rates range between six to fifty percent between two survey rounds. But the intervals between these two rounds differ across country surveys. For comparative purposes they calculate yearly attrition rates for these surveys which range between 1.5 to 23.2 percent (*ibid*:81). An overall panel attrition rate of forty percent in a five year interval between KIDS waves one and two translates into a yearly attrition rate of 13 percent for women of childbearing age in KIDS.[11] This yearly attrition rate is on the upper end of the attrition range documented by Alderman *et al* (2001).

Sample attrition in panel datasets is a major concern for analysts in that it may lead to selective samples and distorted estimation results (*ibid*:80). This is particularly a concern in developing countries, and South Africa in particular, where high levels of mobility or migration exacerbate the problem of attrition (*ibid*:80). In KIDS 1998 specifically, women

---

[10] Alderman *et al* (2001) analyse sample attrition in panel data sets from Bolivia, Nigeria, India, Malaysia and Indonesia.

[11] Following Alderman *et al* (2001:81) the yearly attrition rate is calculated as "$1 - (1- q)^{1/T}$ where q is the overall attrition rate and *T* is the number of years covered by the panel". Given a five year interval between waves one and two of KIDS, T equals 5.

who could not be located or were non-resident are likely to be migrants who have moved to other areas with better employment opportunities. In this case attrition is likely not to be random but selective on variables observed (or unobserved) in the data. If an observed outcome variable of interest affects an individual's propensity for attrition, then attrition is 'non-ignorable' and would bias estimation results (*ibid*:87).

I test whether attrition in the KIDS panel is selective on observables by comparing mean characteristics measured in wave one (1993) for women who were re-interviewed in 1998 versus women who were not re-interviewed (i.e. fell out of the panel in 1998 due to death, migration or because their household was not located in 1998). The T-tests in the last column of Table 2.5 show statistically significant differences in some mean characteristics suggesting that attrition does occur on select observed characteristics. Compared to women re-interviewed in 1998, women who were not re-interviewed were younger (specifically more likely to be aged 20 to 29) and less likely to be married. Being young and single would increase mobility and therefore attrition out of the panel. Women not re-interviewed were more likely to be living in rural areas in 1993 and in slightly poorer households indicated by their reduced access to piped water and electricity. There are neither significant educational differences nor differences in the proportion of women who had experienced at least one live birth by 1993 across the two groups. But when restricting on women who had experienced at least one live birth, fertility was higher among women re-interviewed than women not re-interviewed (see Table 2.6). Therefore fertility estimates using KIDS data would be biased as the birth data appear to be correlated with a woman's propensity for attrition.[12]

[12] It must be noted that I have used a univariate comparison of means to make conclusions about whether attrition is selective on fertility and therefore non-ignorable in a motherhood study. Alderman *et al* (2001) argue that multivariate tests of attrition may yield different conclusions. For example, a probit estimation of the effects of observable variables such as fertility on the probability of attrition may show no significant relationship between fertility and attrition when controlling for other observed characteristics.

Table 2. 5: Mean characteristics of women re-interviewed / not re-interviewed in the second wave of KIDS, African women KwaZulu-Natal

	Re-interviewed		Not re-interviewed		Mean Difference	T-test
Rural	0.719	(0.450)	0.778	(0.416)	-0.059***	-2.8171
Age	26.653	(8.606)	25.685	(7.633)	0.973**	2.4471
Married	0.308	(0.462)	0.214	(0.410)	0.094***	4.4063
Age distribution:						
15 – 19	0.281	(0.500)	0.245	(0.430)	0.036*	1.7050
20 – 24	0.200	(0.400)	0.262	(0.440)	-0.062***	-3.0935
25 – 29	0.142	(0.349)	0.205	(0.404)	-0.063***	-3.5047
30 – 34	0.152	(0.3600	0.128	(0.334)	0.025	1.4554
35 – 39	0.119	(0.324)	0.090	(0.287)	0.029*	1.9457
40 – 44	0.105	(0.307)	0.070	(0.255)	0.035**	2.5690
Education:						
Years of schooling	5.586	(3.111)	5.723	(3.177)	-0.137	-0.9043
No schooling	0.093	(0.290)	0.078	(0.268)	0.015	1.0852
Primary	0.356	(0.479)	0.365	(0.482)	-0.008	-0.3583
Incomplete secondary	0.434	(0.496)	0.423	(0.494)	0.011	0.4602
Matric	0.097	(0.296)	0.116	(0.320)	-0.019	-1.2922
Postmatric	0.021	(0.142)	0.019	(0.136)	0.002	0.2479
Household composition and resources:						
Household size	8.293	(4.088)	8.151	(3.941)	0.142	0.7329
Woman's mother in household	0.520	(0.500)	0.571	(0.495)	-0.051**	-2.1306
Woman's father in household	0.317	(0.466)	0.332	(0.471)	-0.015	-0.6673
Connected to piped water	0.280	(0.449)	0.241	(0.428)	0.039***	1.8541
Connected to electricity	0.352	(0.478)	0.264	(0.441)	0.088***	3.9367
N	1,024		743			
Has had at least one birth	0.656	(0.475)	0.627	(0.483)	0.029	1.2245
N	983		711			

Source: KIDS 1993 and 1998

Notes: The table presents a woman's average characteristics in 1993. Standard errors are in parentheses. Married does not include individuals cohabiting with a partner. Individuals not re-interviewed include individuals who had died since 1993 or were not located in 1998 or were no longer resident in the same household in 1998. The sample includes resident African women aged 15 to 44 in 1993. Results for "has at least one birth" are calculated for a smaller sample of women than other variables because some women had missing birth status information. Married women are identified as women who report having a spouse. Women who are cohabiting with a partner are not identified as married. Mean difference is calculated by subtracting the mean for women not interviewed in 1998 from the mean for women that were interviewed. T-test determines whether the mean difference is significantly different from zero. * 10% significant level; **5% significance level; ***1% significance level.

Table 2. 6: Mean characteristics of women re-interviewed / not re-interviewed in the second wave of KIDS *who had given birth* by 1993, African women KwaZulu-Natal

	Re-interviewed	Not re-interviewed	Mean Difference	T-test
Number of births	3.215 (2.428)	2.587 (1.819)	0.628***	4.6364
N	645	446		

Source: KIDS 1993 and 1998

Notes: The table presents a woman's average characteristics in 1993. Standard errors are in parentheses. Samples include African women in KwaZulu-Natal who have experienced at least one live by 1993. Individuals not re-interviewed include individuals who had died since 1993 or were not located in 1998 or were no longer resident in the same household in 1998. * 10% significant level; **5% significance level; ***1% significance level.

I am not able to make conclusions about attrition on selected labour market outcomes because there were very high levels of non-response on questions related to labour market outcomes among women in KIDS. This is another reason why the KIDS panel cannot be used in a study of motherhood and its economic outcomes.

Measurement error in birth data is also a serious concern in the KIDS. In Table 2.7 I present responses on children ever born for the 1,024 resident women who could be tracked from 1993 to 1998. The results were obtained by differencing a woman's response on number of children ever born in 1998 from her response in 1993. Plausible responses were obtained for a majority of the sample: 47 percent experienced no change in children ever born and 35 percent bore one to four children between 1993 and 1998. But logically impossible responses were obtained for 13 percent of the sample suggesting measurement error in birth data in 1993 or 1998. About 12 percent of the sample "lost" children ever born (i.e. a woman reported fewer numbers of births in her lifetime in 1998 than in 1993); and a further one percent reported unlikely birth entries (in particular eight individuals were reported as having five to ten births between 1993 and 1998 which may be implausible within a five year period).

Problems of missing or incomplete birth status information are also more severe in panel surveys than cross-sectional surveys because women may not respond to birth questions over multiple interviews. It is possible to distinguish zero birth entries from missing birth entries in the KIDS panel. Results in Table 2.7 show that among the 1,024 women about

four percent had missing birth status in 1993. But by 1998, six percent of women had missing birth status in either 1993, 1998 or both years. Excluding women with missing birth status and measurement error in their birth status yielded a final sample of 826 women which could be used in a study of motherhood (representing less than half (826/1,767) of the original 1993 sample).

Table 2. 7: Children ever born (CEB) - transition from 1993 to 1998, African women KwaZulu-Natal

	Frequency	Percent
CEB unchanged	467	45.6
Bore 1 - 4 children	359	35.1
Bore 5 - 10 children	8	0.8
"Lost" CEB	122	11.9
Missing information on CEB in 1993	38	3.7
Missing information on CEB in 1998	27	2.6
Missing information on CEB in both 1993 and 1998	3	0.3
Total	1,024	100

Source: KIDS 1993 and 1998.
Notes: The sample is African women in KwaZulu-Natal that were resident household members in both 1993 and 1998. The women are aged 15 to 44 in 1993 and 20 to 49 in 1998.

2.6 El-Badry hypothesis and adjustment for error in reporting of childless cases

The above analysis of the quality of birth data has highlighted significant levels of missing birth data in both cross-sectional and panel surveys in South Africa. But there is strong evidence to believe that a large proportion of women with missing birth information are actually childless. El-Badry (1961) argues that in developing country surveys a large portion of missing responses should in fact be childless responses. The problem of "zero-errors", where non-mothers are underreported, may arise where the enumerator fails to record a zero entry or childless answer, recording the answer rather as "not-given" or leaving it as a blank response (*ibid*:910-911). This may also arise because the enumerator thinks that asking a question on the number of children ever born is not applicable to women who have not had a birth.

Another related error is "not at-home" error which arises when an enumerator has to ask an individual in the household, who is neither the woman in question nor a knowledgeable

adult, about the number of children ever born to a woman. A lack of full knowledge by the respondent results in the enumerator recording the response in the birth question as "not-given" (*ibid*:910). Results from the OHS 1995 data suggest that not-at-home error is apparent in South African survey data. In Table 2.8 below, not-at-home error is reflected in a higher proportion of missing birth responses when the woman in question is not present during the interview compared to when the woman is present. Among women aged 15 to 49 who were *not* present during the interview, about 24 percent have missing birth entries while this was only 15.2 percent among women who were present during the interview.

El-Badry argues that zero-error is especially apparent when there are a large number of childless women. Younger women are more likely to be childless increasing the likelihood of zero-error reporting among these age cohorts. Not-at-home error is also detected by a higher incidence of non-response on children ever born in the younger ages "because the not-at-home error is in general negatively correlated with the number of children" (*ibid*:912) and younger women have fewer children. In other words, not-at-home error is more likely to occur among younger women compared to older women with more children because their birth status may be less obvious (or more sensitive) to the individual responding on their behalf. El-Badry argues that it is not possible to distinguish between not-at-home error and zero-error in censuses and he therefore uses a combined term "childlessness error" to refer to both errors. Childlessness error is present in the Census 1996 given a higher incidence of non-response among younger age cohorts in South Africa. This was graphically depicted in the downward pattern in Figure 2.1.

Table 2. 8: The proportion of women with missing birth status aged 15 to 49 by whether or not they were present during the household interview

	Present at interview	***Not* present during interview**
Missing birth information	0.152 (0.004)	0.236 (0.007)

Source: OHS 1995.
Notes: The sample above is all women aged 15 to 49. Results are adjusted for weighting and clustering. Standard errors are in parentheses.

El-Badry proposes a method to estimate the number of childlessness errors in each age group so that they can be shifted from the missing to the childless category (*ibid*:915). If childlessness error is more likely to occur when there are a large number of childless cases, then a linear or nearly-linear association exists between the proportion of actually childless cases and the proportion of women whose birth status is missing (*ibid*:915). This positive relationship is evident in the Census 1996 data, reflected in Figure 2.2. The percentage of women in each age cohort with missing birth status is plotted against the percentage of women in each age cohort who are childless. At older age cohorts the percent of women who are childless and the percent of women who have missing birth data are both low, while at younger age cohorts these percentages are higher.

Mathematically the linear relationship between the percent of women with missing birth data and the percent of women who are childless can be represented as follows:

$$Z_i = aN_i^* + b \qquad (a)$$

Z_i is the reported proportion of women with missing birth status which is made up of two components aN_i^* and b. Component 'b' is the *true* proportion of women whose birth status is missing while aN_i^* is the proportion of women whose birth status was *erroneously* reported as missing when in fact they are childless (Moultrie and Timaeus 2002:69, El-Badry 1961:915). The coefficient 'a' reflects the degree of incidence of the childlessness error i.e. the true proportion of women with no children in age group i who were incorrectly recorded as having missing birth status. N_i^* is the true proportion of women in each age group *i* that are childless. But the true proportion of childless women, N_i^*, is higher than the reported proportion of childless women. I denote the reported proportion of childless women as N_i which is equal to the true proportion of childless women less the proportion of childless women who are erroneously reported as missing:

$$N_i = (1-a)N_i^* \qquad (b)$$

If this equation is rearranged and substituted into (a) then

$$Z_i = \frac{a}{(1-a)} N_i + b$$

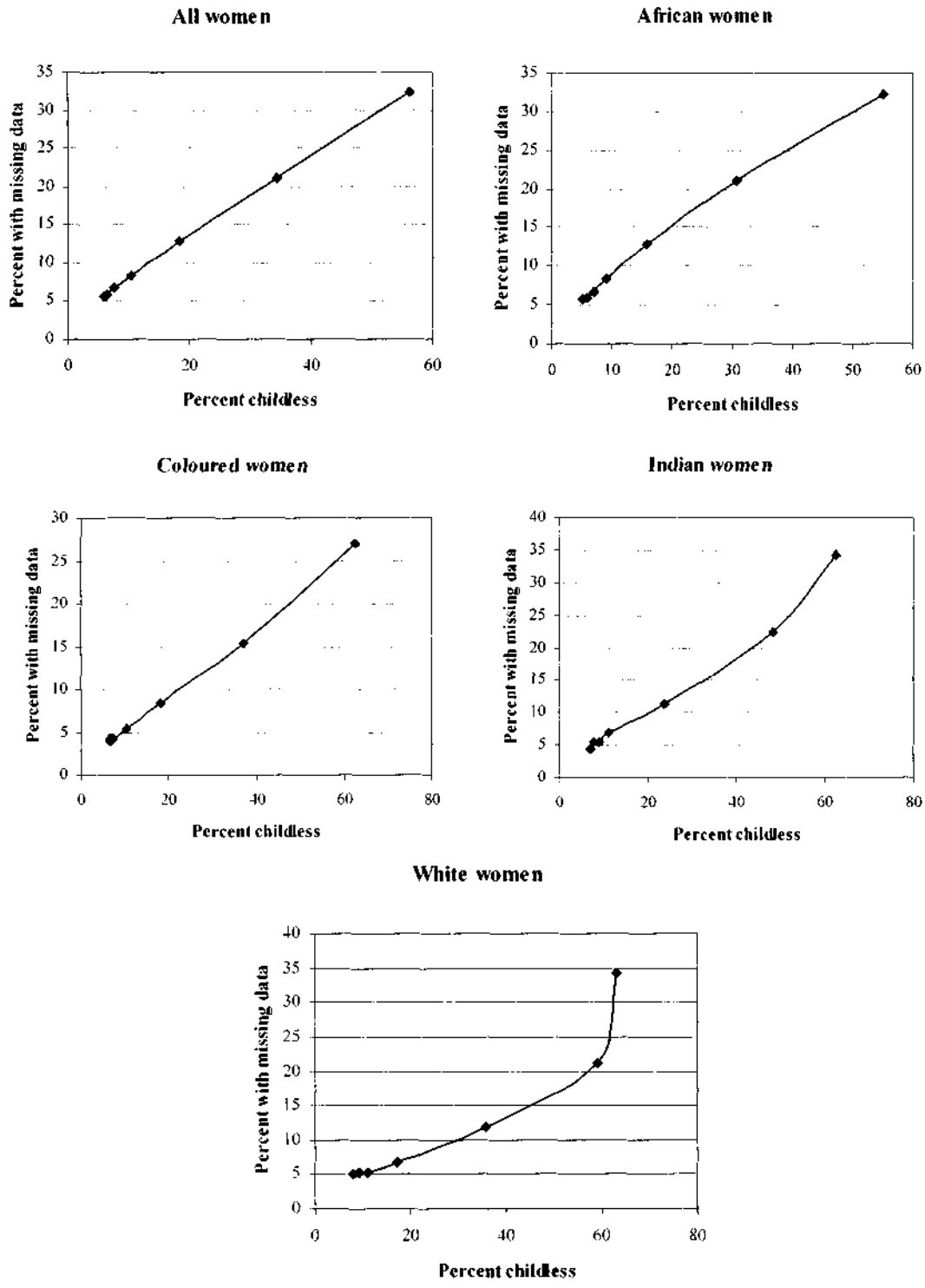
It is possible to determine the value of b i.e. the *true* percent of women with missing birth data. It is simply the constant term estimated when fitting a linear regression of the reported percent of women with missing birth data, Z_i on the reported percent of childless women in each age category, N_i. Having determined b, Z_i - b then represents the proportion of women with missing birth status that should be adjusted to childless status. The true proportion of childless women, N_i* is therefore represented as follows:

$$N_i^* = N_i + (Z_i - b)$$

Following Moultrie and Timmaues (2002), I perform an El-Badry adjustment on the Census 1996 data to generate an estimate of the proportion of women with missing birth data that are in fact childless. I perform an ordinary least squares (OLS) regression of the percentage of women with missing birth data on the percentage of childless women at every age from 20 to 49. The results of the regression are presented in columns indicated four and five in Table 2.10.

The regression result of interest is the constant, b, which reflects the 'true' percent of women whose birth data are missing. Compared to column one in Table 2.9 which reflects the percent and counts of women with *reported* missing birth status, the estimated results for 'true' missing birth status in column five and six (Table 2.10) are much lower. For example, among all women the estimated 'true' percent with missing birth data is 2.6 percent as opposed to a reported figure of 10.2 percent. Knowing b, I calculate the percentage of women erroneously reported as having missing birth status but who are in fact childless ie. $Z_i - b$ (see column seven of Table 2.10). Multiplying $Z_i - b$ by the total number of women in each relevant group (column three in Table 2.9) yields the number of women with missing birth status that are shifted into the childless category, reflected in column eight (El-Badry 1961: 916).

Figure 2. 2: Missing birth status and childless status, by age category



Source: Census 1996.
Notes: The data are unweighted. Age categories are in descending order: 45-49, 40-44, 35–39, 30-34, 25-29, 20-24 and 15-19.

Table 2. 9: Reported results of the percent of women aged 20 to 49 having missing birth status and childless status, 1996

Population group	1. Reported % i.e. Z_i, and counts of women with missing birth status	2. Reported % i.e. N_i and counts of childless women	3. Counts of *all* women aged 20 to 49
African	10.979 708,357	14.983 982,916	6,560,305
Coloured	6.742 55,253	16.210 134,116	827,359
Indian	9.562 24,037	19.876 504,13	253,638
White	7.873 77,355	24.591 244,863	995,738
All	10.221 876,342	16.371 1,425,365	8,706,815

Source: Census 1996.
Notes: Sample is of all women aged 20 to 49. Standard errors are in parentheses. Data are not weighted.

Table 2. 10: Results of the El-Badry adjustment to Census 1996 data

Population group	4. Coefficient on the reported % of childless women i.e. a/(1-a)	5. *True* % of women with missing birth status, i.e. b (constant)	6. The *true* counts of women with missing birth status	7. % and counts of all women to be shifted from missing to childless category i.e. Z_i - b	8. True % i.e. N_i*, and counts of childless women
African	0.604 *** (0.011)	3.142*** (0.166)	206,125	7.837 502,232	22.820 1,485,148
Coloured	0.365 *** (0.008)	1.719*** (0.156)	14,222	5.023 41,031	21.233 175,147
Indian	0.432*** (0.019)	1.698*** (0.433)	4307	7.864 19,730	27.740 70,143
White	0.314*** (0.015)	1.911*** (0.461)	19,029	5.962 58,326	30.553 303,189
All	0.543*** (0.005)	2.564*** (0.094)	223,243	7.657 653,099	24.028 2,078,464

Source: Census 1996.
Notes: Sample is of all women aged 20 to 49. Standard errors are in parentheses. *** 1% significance level. The data are not weighted.

Comparing the reported counts of childless women in column two with the El-Badry adjusted counts of childless women in column eight, there is a marked increase in childless women. The count of childless women increases by approximately 653,000. The remaining counts of women with 'true' missing birth data suggest an upper limit by which mothers may be underestimated. Supposing all the remaining counts of women with missing birth status were mothers, then women aged 20 to 49 who had a live birth would be underestimated by less than 223,000 (or 2.6 percent of all women aged 20 to 49). This indicates that the extent to which mothers are underestimated in household surveys is smaller than high levels of recorded missing birth data may suggest. Unfortunately the El-Badry adjustment could not be successfully applied to the PSLSD 1993 and the OHS 1995 data to verify these results because of their smaller sample sizes.

An important qualification must be made when analysing the results of the El-Badry adjustment. The hypothesis rests on a linear association between the number of childless cases in each age cohort and the number of missing responses. This linear association is based on the primary assumption that there is a greater likelihood of zero or not-at-home reporting errors when the number of childless cases is greater – particularly in younger age cohorts. However higher levels of birth non-response in younger age cohorts may also be the result of secrecy surrounding sex and reproduction. The El-Badry hypothesis does not account for this and the results of the adjustment may be biased in younger age cohorts. Young women with missing birth data shifted into the childless category may have actually been true non-respondents choosing not to disclose fertility information to the enumerator while in the presence of adults or parents. In this case, the El-Badry adjustment overestimates the number of childless women and underestimates the number of true non-respondents.

Conclusion

This chapter identified two sets of methods for identifying mothers using household surveys and examined the data available to match women to their children. A clear comparison was presented of the key aspects of these methods explaining their similarities, differences and limitations. The first set of methods (method one and two) limit mothers to women who have children who are biologically their own. Method one identifies mothers as women who have had at least one live birth while method two further restricts mothers to women who have had a live birth *and* the child is still alive. This first set of methods relies on data collected in birth modules. However the quality of birth data in many cross-sectional and panel surveys is compromised by high levels of non-response on children ever born.

The poor quality of birth data in South African surveys was highlighted by analysing incomplete or missing birth status in the PSLSD 1993, the OHS 1995 and the Census 1996. In all three surveys non-response was particularly high among younger age cohorts but specifically in the OHS 1995 and the Census 1996. An El-Badry adjustment to Census 1996 data suggested that a key reason for high levels of non-response is enumerator errors in reporting where non-mothers or childless women are incorrectly recorded as having missing birth status. The majority of women with missing birth data may actually be childless: over 650,000 women with reported missing birth data or an *additional* eight percent (=653,099/8,706,815) of all women aged 20 to 49 were shifted into the childless category. After the adjustment, the remaining *true* number of women with missing birth status was only about 223,000 which is less than 2.6 percent of all women in South Africa. This provides evidence that the training of enumerators to fill out accurately birth module information could considerably reduce problems of non-response in birth data.

The results of the El-Badry adjustment on the Census 1996 data also indicate that the extent to which mothers may be underestimated is lower than the high levels of missing birth data may suggest because the majority of women with missing birth status are actually non-mothers. Although conclusions of the El-Badry adjustment mitigate concerns about the

quality of birth data in cross-sectional surveys, the quality of birth data remains particularly compromised in panel datasets. This was illustrated by analysing birth data in the KwaZulu-Natal Income Dynamics Study (KIDS) in which panel attrition, measurement error as well as missing birth data render it unusable for a study of motherhood and its economic costs. When tracking the birth status of women using waves one and two of KIDS, sample attrition of African women is over forty percent between the five year period or 13 percent per annum. This attrition rate is high relative to other developing country panel surveys. Furthermore attrition occurs on select observed characteristics so that the remaining sample of mothers who do not fall out of the panel is likely to be a non-random sample. Specifically they have higher fertility rates than women who do not fall out of the sample.

A key reason for high levels of sample attrition is that KIDS does not collect birth status information on women who become non-resident household members. Future panel surveys should aim to track and interview women who become non-residents on their birth status. This will considerably reduce sample attrition in a motherhood study and allay concerns that selection on observable characteristics will bias estimated results; but this may also be very difficult and expensive to execute, particularly in nationally representative surveys. A cheaper but less reliable alternative would be to collect birth status information on non-resident women from a knowledgeable resident household member. These solutions, however, will not solve problems of measurement error in birth data which is collected for the same women in repeated surveys. Measurement error on children ever born (and missing birth status) further reduced the number of women whose birth status could be tracked in KIDS from 1993 to 1998; birth status for less than half (826/1,767) of the original 1993 sample of African women aged 20 to 44 could be identified in 1998.

In order to avoid issues of birth data quality in identifying mothers, a second set of methods (methods three and four) were identified. These methods use household relationship questions rather than birth data to match women to their children, and can be used to identify both biological *and* non-biological mothers. This is useful in the South African context of extended family structure and high levels of child fosterage where childcare is

often provided by women who are not the child's biological mother, for example grandmothers or aunts. However a necessary requirement to identify mothers using methods three and four is that they are resident in the same household as their children. In the context of female migration and inter-household mobility, a large proportion of mothers in South Africa may not to be co-resident with their children. Methods three and four will substantially underestimate the true number of mothers in South Africa. In addition, method four, which uses questions on kin relationships to the household head, is subject to misclassification.

This chapter has highlighted that each of the four identification methods has some limitation in providing accurate measures of mothers in South Africa. Consequently, I suggest that a set of motherhood measures from different methods be used. Considered in conjunction, their estimates will provide a more comprehensive and interesting picture of motherhood. In particular, it becomes possible to explore mother-child residency patterns in South Africa.

Appendix 2A: Questions on who a person's mother is in the household

PSLSD 1993:

Q.9. If the mother of _____ lives here, write the mother's code. If absent, code 99; if deceased code 88

OHS 1995:

2.3 Are the parents of ... still alive?	Father	Mother
1 = Yes 2 = No	1 2	1 2
2.4 Give respondent numbers of parents of ... if they are part of the household.	Father	Mother

GHS 2002:

1.4.a	**Is ……'s mother still alive?** 1 = YES 2 = NO } → *Go to Q 1.5.a* 3 = Don't know
1.4.b	**Is ……'s mother part of the household?** 1 = YES 2 = NO → *Go to Q 1.5.a*
1.4.c	**Which person is ……'s mother?** *Give person number*

Appendix 2B: Questions asked on a person's relationship to the household head

PSLSD 1993:

3. **What is _____ 's relationship to house-hold head?**
01=Resident Head 02=Absent Head 03=Wife or husband or partner 04=Son or daughter 05=Father or mother 06=Grandchild 07=Grandparent 08=Mother- or father-in-law 09=Son- or daughter-in-law 10=Brother- or sister-in-law 11=Aunt or uncle 12=Sister or brother 13=Niece or nephew 14=Cousin 15=Great-grandparent 16=Household help (or relative of) 17=Lodger or relative of lodgers 18=Other family 19=Other non-family

1995 OHS:

2.2 Relationship to head/acting head of household 1 = HEAD (or reference person) 2 = Husband/wife/partner 3 = Son/daughter 4 = Parent 5 = Grandparent 6 = Grandchild 7 = Brother/sister 8 = Other relative (e.g. in-laws or aunt/uncle) 9 = Non-related person

GHS 2002:

1.1	**What is ……'s relationship to the head of the household?** (*I.e. to the person in column 1)* 1 = *Mark the head/acting head* 2 = HUSBAND/WIFE/PARTNER 3 = SON/DAUGHTER/STEPCHILD/ADOPTED CHILD 4 = BROTHER/SISTER 5 = FATHER/MOTHER 6 = GRANDPARENT/GREAT GRANDPARENT 7 = GRANDCHILD/GREAT GRANDCHILD 8 = OTHER RELATIVE (E.G. IN-LAWS OR AUNT/UNCLE) 9 = NON-RELATED PERSONS

Chapter Three:

Measuring mothers in South Africa

Introduction

There has been renewed interest in examining fertility in South Africa as more birth data have been made available since the 1990's (Moultrie and Timaeus 2002). Focus has centred on identifying fertility rates, trends in fertility rates as well as the determinants of fertility. Studies have has also considered how fertility rates vary by demographic and socio-economic characteristics (Palamuleni *et al* 2007; Moultrie and Timaeus 2001; Udjo 2000; Mencarini 1999; Mostert and Hofmeyr 1988). There is little research which has measured mothers in South Africa or considered identification problems in measuring the extent of motherhood.

In this chapter, I apply the identification methods outlined in the previous chapter to investigate *how many* and what *proportion* of women are mothers in South Africa using cross-sectional household surveys. I consider two sets of identification methods to measure the extent of motherhood. In the first set, I use birth status information to generate measures of biological mothers, regardless of whether or not they co-reside with their children. In the second set, I use questions on household relationships to measure biological *and* non-biological mothers. But a necessary requirement to identify mothers using household relationship questions is that mothers are co-resident with at least one of their children. Therefore the second set of measures imposes a co-residency restriction on motherhood. By comparing the two sets of measures, I highlight that motherhood measures differ considerably by whether they include 'absent' (or not co-resident) mothers and non-biological mothers.

Political and social changes in post-apartheid South Africa, such as the removal of controls on urbanisation among Africans and the increasing AIDS epidemic may have impacted upon household living arrangements, including the co-residency of mothers with their children. Comparing the first and second set of motherhood measures over time, this

chapter reveals that an increasing proportion of mothers were not living with any of their children in 2002 compared to 1993. I am able to show that measures that place a co-residency restriction on motherhood have increasingly underestimated the true extent of motherhood in South Africa.

The chapter also highlights that when birth status information is used to identify mothers, motherhood measures are often inconsistent and very 'noisy' across different household surveys. I therefore interrogate the robustness of motherhood measures derived from birth status information.

Section 3.1 considers sample descriptions and data used in the chapter. Section 3.2 presents the first set of motherhood measures derived from birth status information and section 3.3 presents the second set of measures that rely on household relationship questions to match women to their children. Section 3.3 also compares these two sets of measures, highlighting that a sizeable sample of all mothers (and specifically African mothers) are not co-resident with their children. Finally in section 3.4, I show that when household relationship questions are used to generate measures of co-resident mothers, maternal relationship questions result in more accurate measures than questions on relationships to the household head.

3.1 Data and definitions

I use the 1993 Project for Statistics on Living Standards and Development (PSLSD), the October Household Survey (OHS) 1995, the ten percent sample of the Population Census 1996 and the General Household Survey (GHS) 2002 to measure mothers. I also consider measures from the OHS 1997 and the ten percent sample of the Population Census 2001. Comparing measures across the different surveys, I am able to determine which of the following datasets provide reasonably consistent measures of mothers.

Throughout the analysis the sample is restricted to all women aged 20 to 49. The sample is also disaggregated by race. I consider two sets of motherhood measures. The first set uses

birth status information to determine what proportion and number of women has *biological* children, regardless of whether these women co-reside with their children. Here I use identification methods one and two outlined in Chapter Two to match women to their children using birth data. Method one identifies a woman as a mother if she has had a least one live birth, regardless of whether the child is still alive. Method two further restricts mothers to women who have had at least one live birth *and* at least one of these children is still alive.

The second set of measures relies on household relationship questions to generate proportions and counts of women who are *co-resident* with at least one of their biological *or* non-biological children. Here I use methods three and four, outlined in Chapter Two, to identify mothers. Method three identifies mothers using a maternal relationship question while method four links women to children using a question on relationships to the household head. In contrast to the first set of measures, the second set imposes a co-residency requirement on motherhood. However in both sets of measures, motherhood is not restricted on the age of children belonging to a woman.

Throughout this analysis I provide both counts of mothers aged 20 to 49 in South Africa and proportions of women aged 20 to 49 who are mothers. Proportions are calculated as follows:

$$\text{proportion of women who are mothers} = \frac{\text{number of mothers aged 20 to 49}}{\text{all women aged 20 to 49}}$$

This is a straightforward calculation if mothers are identified using method three and four because they rely on household relationship questions that are not subject to high levels of non-response. However, when using methods one and two, this calculation and its comparison across different surveys, is complicated by missing or incomplete birth data in household surveys. Across national surveys which ask information on children ever born, there are variations in the coding of true non-respondents and childless women. In the PSLSD 1993, the OHS 1995 and the Census 1996, the birth module question on number of

children ever born allowed a specific response for no children. A positive entry could be made for mothers who reported having at least one child, a zero entry could be made for childless women (or non-mothers) while a missing response was given for true non-respondents. In the GHS 2002 and the OHS 1997, however, a zero entry was *not* allowed for childless women. True non-respondents and childless women were both coded as non-respondents and therefore could not be distinguished from one another.

A question arises about what to do with missing birth status data across the different surveys. The El-Badry adjustment to the Population Census 1996 data indicated that most non-responses are likely to represent non-mothers (see Chapter Two for more details). If all women with missing birth status data are dropped from the sample used in the calculation then measures of the true proportion of women who are mothers would be *overestimated* using the PSLSD 1993, the OHS 1995 and the Census 1996 as well as the OHS 1997 and the GHS 2002. The reason is that a significant proportion of non-mothers would be excluded from the denominator of the calculation. In fact, in the OHS 1997 and the GHS 2002 all non-mothers would be excluded from the denominator because all non-mothers are coded as missing birth entries; therefore, the proportion of women who are mothers would equal one.

Another problem with dropping missing data from the calculation is that the samples from the aforementioned surveys will not be consistent with the ten percent sample release of the Population Census 2001. In the Census 2001, there are no missing birth data entries because birth data were imputed for non-respondents before the public release of the data. Women either have a positive birth entry or a zero birth entry.

In order to obtain motherhood measures that are comparable across the different surveys, I convert women with missing birth status to 'zeros' or non-mothers in the PSLSD 1993, the OHS 1995, the OHS 1997, the Census 1996 and the GHS 2002 to match coding in the Census 2001. The proportion of women who are mothers is therefore calculated as a total of *all* women aged 20 to 49, regardless of whether a woman's birth status was reported.

There will be a consistent problem across the PSLSD 1993, the OHS 1995, the OHS 1997, the Census 1996 and the GHS 2002 of potentially *underestimating* the true proportion of women who are mothers when converting all missing birth responses to zeroes. The reason is that included in the zero responses will be a small proportion of women who are actually mothers and did not report their birth status information. However the discussion in the following section suggests that the Census 2001 may *overestimate* the true proportion of women who are mothers as a result of incorrect imputation of birth status information.

3.2 The first set of measures: using birth status information

Birth status information can be used to identify biological mothers. Birth modules typically ask a question on the number of children ever born alive to a woman. I present measures derived from this question for a national sample of women aged 20 to 49 and then disaggregated by race groups. I also investigate whether further restricting these measures to exclude women whose every child has died since birth makes a difference to measures of biological mothers.

A key finding in this section is that measures of biological mothers are often not consistent across different household surveys. I therefore evaluate the robustness of motherhood measures, by comparing measures across different household surveys.

3.2.1 Measures of a national sample of biological mothers

Table 3.1 shows the proportion and counts of women aged 20 to 49 in South Africa who have had at least one live birth (i.e. are biological mothers), regardless of whether these children are still alive (i.e. method one). It also presents total population counts of women aged 20 to 49. The table distinguishes between surveys according to the consistency of their motherhood measures (I will comment on the consistency of measures later in this section). I first discuss more robust measures from the PSLSD 1993, the Census 1996 and the GHS 2002 that are presented in the shaded columns.

Table 3. 1: Proportions and counts of all women aged 20 to 49 who have had at least one live birth (method one)

	PSLSD 1993	OHS 1995	Census 1996	OHS 1997	Census 2001	GHS 2002
Proportion	0.767 (0.007)	0.710 (0.005)	0.733 (0.000)	0.646 (0.004)	0.815 (0.000)	0.768 (0.004)
Counts of mothers	6,314,914 (154,327)	6,142,656 (1,276,440)	6,379,713 (7,760)	5,898,094 (67,116)	7,567,592 (5,853)	7,670,496 (102,847)
Counts of all women 20 to 49	8,235,384 (202,375)	8,650,612 (1,770,653)	8,706,815 (8,754)	9,131,915 (88,321)	9,280,967 (9,326)	9,992,795 (130,541)

Notes: Standard errors are in parentheses. Estimates are obtained by accounting for weighting and clustering in each survey's sample design. The sample includes all women aged 20 to 49. The PSLSD 1993 captured information for both resident and non-resident women but the sample above is limited to resident women only, as birth status was only asked of resident women in the household. No estimates are presented for women in the Census 1996 and the GHS 2002 with race reported as "other" or "missing". Counts and standard errors on counts are rounded up to the nearest unit.

In 1993, 77 percent of a national sample of women aged 20 to 49 were biological mothers, representing an absolute count of about 6.3 million women. However 4.5 percent of women in the PSLSD 1993 had missing birth status (see Chapter Two for more details). All of these women with missing birth status have been coded as non-mothers (or zeroes) in obtaining motherhood measures; however if some of these women are in fact mothers then more than 77 percent of women were mothers in 1993.

In 1996, there was slight increase in the number of women who were mothers to about 6.4 million, but the increase in the total population of women aged 20 to 49 was comparatively larger. The proportion of women who were mothers therefore fell to 73 percent. Higher levels of missing birth data of 10.2 percent in the Census 1996 compared to 4.5 percent in the PSLSD 1993, suggest that underreporting of mothers is more severe in 1996 compared to 1993. However results from the El-Badry adjustment to the Census 1996 data suggest that the *true* proportion of women with missing birth data was only 2.6 percent after shifting truly childless women from the missing data category into the childless category (see Chapter Two for more details). If 2.6 percent is an upper limit to the amount by which

mothers are underestimated, then no more than 76 percent of women (or an additional 223,000 women) aged 20 to 49 had experienced a live birth in 1996.

By 2002, absolute counts of mothers increased to approximately 7.7 million. Compared to 1996, the proportion of all women reported as having had at least one live birth increased to 77 percent. This figure is likely to represent a lower band for mothers because all non-responses, which may include mothers, were coded as non-mothers (zeroes). Unfortunately the extent of this non-response cannot be identified.

The motherhood measures in the shaded columns of Table 3.1 are generally robust across the 1993 PSLSD, the Census 1996 and the GHS 2002. Together they suggest that among a national sample of women aged 20 to 49, the proportion that were biological mothers was about 77 percent in 1993, fell to 73 percent in 1996 and then increased again to about 77 percent in 2002. Although these measures are possibly underestimated given that some mothers may be coded as non-mothers (or zeroes), the El-Badry adjustment to the Census 1996 data suggests that this underestimation is less than 2.6 percent.

I now consider motherhood measures from the OHS 1995, the OHS 1997 and the Census 2001, presented in columns that are not shaded in Table 3.1. These measures are inconsistent both with each other and with measures from the surveys in the shaded columns. Motherhood trends are very noisy across the OHS 1995, the OHS 1997 and the Census 2001. The proportion of women who are biological mothers falls by six percentage points from 1995 to 1997 and then rebounds by about 12 percentage points from 1997 to 2001. These dramatic fluctuations are unlikely to represent true fluctuations in the proportion of women who are mothers. Furthermore motherhood measures from the OHS 1995, the OHS 1997 and the Census 2001 deviate considerably from the measures derived from the PSLSD 1993, the Census 1996 and the GHS 2002 (represented in the shaded columns of Table 3.1).

Measures of biological mothers from the OHS 1995 and particularly the OHS 1997 underestimate the true proportion of women who are mothers when compared with other

surveys. It is unreasonable to expect such dramatic falls in the true proportion of all women who were mothers from 1993 to 1995 and again from 1996 to 1997. By contrast, the Census 2001 measures overestimate the extent of motherhood relative to all other surveys. It is however not clear why these surveys generate inconsistent motherhood measures. Problems of missing birth status data and inaccurate imputation of birth status data may offer possible explanations for inconsistent measures. In the OHS 1995 for example, motherhood measures may be compromised by high levels of missing birth data. Among women aged 20 to 49, non-response on children ever born in the OHS 1995 was about 12 percent.[13] A proportion of women coded as non-respondents may actually be mothers. However it is not possible to determine the extent of missing birth data in the OHS 1997 and the Census 2001.

A key explanation for the inconsistency of motherhood measures generated from the Census 2001 is the inaccurate imputation of missing birth status information. Moultrie and Dorrington (2004) analyse the imputation procedures used by Statistics South Africa to fill in blank or missing birth status entries in the Census 2001. They note that edit rules used to fill in these blank entries did not take into consideration the findings by El-Badry (1961) that a large proportion of women with missing birth status are in fact childless women, especially among the younger age cohorts (Moultrie and Dorrington 2004:9). Hotdeck imputations for missing birth status have the effect of imputing children to women who are in fact childless (ibid:9). Consequently, there is reason to believe that the number of mothers identified in the Census 2001 is biased upwards.

3.2.2 Measures of biological mothers disaggregated by race

I now consider variations in measures of biological mothers across different race groups. Table 3.2 shows how many and what proportion of women has had at least one live birth, disaggregated by race group.

[13] Standard errors on the measures of mothers using the OHS 1995 are also very much larger than in other surveys. Therefore the OHS 1995 estimates lack precision.

Table 3. 2: Proportions and counts of women aged 20 to 49 who have had at least one live birth, by race (method one)

	PSLSD 1993	OHS 1995	Census 1996	OHS 1997	Census 2001	GHS 2002
African						
Proportion	0.797 (0.006)	0.703 (0.005)	0.739 (0.001)	0.652 (0.005)	0.825 (0.005)	0.776 (0.004)
Counts of mothers	4,808,625 (189,678)	4,575,925 (889,509)	4,847,709 (6,896)	4,527,035 (71,204)	6,085,807 (5,980)	5,956,716 (99,704)
Counts of all women 20 to 49	6,033,330 (239,066)	6,513,306 (1,248,024)	6,560,305 (7,826)	6,947,324 (101,389)	7,375,170 (5,842)	7,672,895 (126,165)
Coloured						
Proportion	0.753 (0.018)	0.746 (0.008)	0.770 (0.002)	0.591 (0.014)	0.830 (0.001)	0.809 (0.010)
Counts of mothers	556,234 (101,068)	630,138 (168,345)	636,927 (2,565)	519,518 (29,670)	699,277 (2,744)	772,888 (41,263)
Counts of all women 20 to 49	738,745 (130,845)	844,967 (224,852)	827,359 (2,894)	878,347 (46,090)	842,064 (2,990)	955,744 (49,175)
Indian						
Proportion	0.709 (0.037)	0.721 (0.016)	0.704 (0.003)	0.631 (0.024)	0.722 (0.003)	0.726 (0.023)
Counts of mothers	182,987 (54,455)	184,337 (33,078)	178,668 (1,361)	164,088 (19,732)	172,839 (1,392)	199,911 (21,583)
Counts of all women 20 to 49	258,049 (75,837)	255,722 (44,562)	253,638 (1,619)	259,954 (30,078)	239,234 (1,631)	275,348 (28,012)
White						
Proportion	0.636 (0.024)	0.726 (0.012)	0.674 (0.002)	0.657 (0.014)	0.739 (0.002)	0.682 (0.016)
Counts of mothers	767,068 (111,247)	752,256 (260,940)	671,324 (2,702)	687,453 (41,655)	609,669 (2,790)	736,998 (42,017)
Counts of all women 20 to 49	1,205,260 (176,259)	1,036,616 (354,448)	995,738 (3,270)	1,046,290 (59,345)	824,499 (3,191)	1,081,217 (60,296)

Notes: Standard errors are in parentheses. Estimates are obtained by accounting for weighting and clustering in each survey's sample design. The sample includes women aged 20 to 49. The PSLSD 1993 captured information for both resident and non-resident women but the sample above is limited to resident women only, as birth status was only asked of resident women in the household. No estimates are presented for women in the Census 1996 and the GHS 2002 with race reported as "other" or "missing". Counts, and standard errors on counts, are rounded up to the nearest unit.

Evident from the table is that the extent of motherhood is highest among Africans, followed by Coloureds, Indians and lowest among Whites in 1993. This mirrors patterns reported in South African fertility literature in which total fertility rates follow a similar pattern by race composition (Moultrie and Timaeus 2002; Mencarini 1999). However the Census 1996 and the GHS 2002 results suggest that the extent of motherhood is highest among Coloureds followed by Africans.[14]

The consistency of motherhood measures across different surveys varies by race group. Among African women aged 20 to 49, motherhood measures are relatively consistent across the PSLSD 1993, the Census 1996 and the GHS 2002 (represented in the shaded columns of Table 3.2). In 1993, about 80 percent of all African women aged 20 to 49 were biological mothers but this percentage fell to 74 percent in 1996 and increased to 78 percent in 2002. The general decline in the proportion of women who are mothers is consistent with studies on fertility in South Africa that report a general decline in fertility rates among African women since 1993 (Moultrie and Timaeus 2002). The measures of African mothers from the OHS 1995, the OHS 1997 and the Census 2001, however, are inconsistent both with estimates in the shaded columns and with declining fertility rates among African women. This suggests that the OHS 1995, the OHS 1997 and the Census 2001 provide unreliable measures of African mothers.

With the exception of the OHS 1997, motherhood measures among Indians appear robust across all surveys in Table 3.2. About 71 to 73 percent of all Indian women aged 20 to 49 were biological mothers over the period 1993 to 2002. By contrast, measures of Coloured and White mothers vary considerably across the six different surveys. It is difficult to draw conclusions about a plausible estimate range for the proportion of White and Coloured women who are mothers. Among Coloured women, however, there appears to be a general

[14] This contrasts with Moultrie and Timaues's (2002) finding that total fertility rates are highest among Africans in 1996 followed by Coloureds using Census data. In reconciling these differences, it is possible that relative to African women, fewer children are born to each Coloured mother but a larger proportion of all Coloured women have at least one live birth.

upward trend in the proportion of women who are mothers from 1993 to 2002 (if the OHS 1997 results are ignored).

3.2.3 Excluding women whose every child has died since birth

Because very few mothers are likely to experience the death of *all* their children, measures of biological mothers are not expected to differ considerably by whether they include or exclude mothers with no surviving children. The results presented in Table 3.3 below are consistent with this expectation. This table presents measures of the proportion of women aged 20 to 49 who are biological mothers *and have at least one living child* (method two). For comparative purposes I replicate results from Tables 3.1 and 3.2 which did not restrict biological mothers to women with living children. Again I shade measures from the PSLSD 1993, the Census 1996 and the GHS 2002 to distinguish them from less robust motherhood measures generated from the OHS 1995 and the OHS 1997.

The table suggests that motherhood measures are not very sensitive to the exclusion of women whose every child has died since birth. Within each cell of Table 3.3 the percentage point variation across measures that include and exclude these women is generally less than 1.5 percent. The results in the shaded columns indicate that trends in the extent of motherhood are also relatively unchanged by this exclusion. In 1993 about 75 percent of a national sample of women aged 20 to 49 were identified as having at least one biological child who is still alive. This figure decreased slightly to 72 percent in 1995 and increased to 76 percent in 2002.

The small variation across measures of biological mothers that exclude or include women whose every child has died since birth is expected because few women experience the death of all their children. However reporting errors in birth status information may also explain this finding. First, child deaths after birth may be underreported in household surveys. In this case, motherhood measures which exclude women whose every child has died since birth will be upwardly biased toward measures that include these women. Another commonly sighted problem in the reporting of birth data is that women report total live births as the total number of children still alive (Moultrie and Dorrington 2004). If this

error is made by mothers with no surviving children, they will report total live births as zero. In this case, measures of mothers that include women whose every child has died since birth will be biased downward toward measures which exclude these women.

Table 3. 3: Including / excluding women whose every child has died since birth from motherhood measures (method one vs. method two), women aged 20 to 49

	PSLSD 1993	OHS 1995	Census 1996	OHS 1997	GHS 2002
			All		
Exclude	0.753 (0.007)	0.706 (0.005)	0.719 (0.001)	0.643 (0.004)	0.762 (0.004)
Include	0.767 (0.007)	0.710 (0.005)	0.733 (0.000)	0.646 (0.004)	0.768 (0.004)
			African		
Exclude	0.780 (0.006)	0.698 (0.005)	0.724 (0.000)	0.648 (0.004)	0.770 (0.004)
Include	0.797 (0.006)	0.703 (0.005)	0.739 (0.001)	0.652 (0.005)	0.776 (0.004)
			Coloured		
Exclude	0.751 (0.017)	0.743 (0.008)	0.759 (0.002)	0.589 (0.014)	0.806 (0.010)
Include	0.753 (0.018)	0.746 (0.008)	0.770 (0.002)	0.591 (0.014)	0.809 (0.010)
			Indian		
Exclude	0.709 (0.037)	0.721 (0.016)	0.698 (0.003)	0.628 (0.024)	0.725 (0.023)
Include	0.709 (0.037)	0.721 (0.016)	0.704 (0.003)	0.631 (0.024)	0.726 (0.023)
			White		
Exclude	0.631 (0.024)	0.725 (0.011)	0.667 (0.002)	0.656 (0.014)	0.681 (0.016)
Include	0.636 (0.024)	0.726 (0.012)	0.674 (0.002)	0.657 (0.014)	0.682 (0.016)

Notes: The first line of measures in each cell exclude women whose every child has died since birth from the the sample of mothers while the second line of measures includes these women (and is replicated from Tables 3.1 and 3.2). Standard errors are in parentheses. Estimates are obtained by accounting for weighting and clustering in each survey's sample design. The sample includes women aged 20 to 49. The PSLSD 1993 captured information for both resident and non-resident women but the sample above is limited to resident women only as birth status was only asked of resident household members. No estimates are presented for women in the Census 1996 and the GHS 2002 with race reported as "other" or "missing".

3.3 The second set of results: using relationship questions

In the first set of measures which used birth information status to identify mothers, motherhood was restricted to women with biological children, regardless of co-residency status with children. I now consider a second set of measures derived from household relationship questions.

A key aspect of using household relationship questions to measure mothers is that they can also identify non-biological mothers. In the context of high levels of child fosterage in South Africa, household relationship questions offer an advantage over birth information by identifying non-biological mothers. However a necessary requirement to match women to their children using household relationship questions is that mothers are co-resident in the same household as at least one of their children. Therefore this second set of measures restricts mothers to women who are co-resident with at least one of their children.

The discussion in Chapter Two indicated that there are two household relationship questions which can be used to link women to their children. A maternal relationship question asks an individual to report the person number of his or her mother if she is alive and resident in the same household. This question is available in the PSLSD 1993, the OHS 1995 and the GHS 2002. Another possible method to identify mothers uses a question on relationships to the household head which is available in the 1993 PSLSD, the two population Censuses and all the OHS and GHS surveys (see Chapter Two for more details).

In the first two rows of Table 3.4 I present motherhood measures using the maternal relationship question (identified as method three in Chapter Two) and the question on relationships to the household head (identified as method four in Chapter Two). Here mothers are identified as biological or non-biological mothers who are co-resident with at least one child. The third row reproduces measures in Table 3.3 derived from birth status information. In this case, a mother is identified as a woman with at least one biological child who is still alive, regardless of co-residency status.

Despite capturing non-biological mothers, or women who have assumed care for children who are not their own, the results in Table 3.4 show that measures based on household relationship questions are much smaller than measures based on birth status information. The percentage of all women aged 20 to 49 who are mothers is about six to 12 percentage points lower when derived from household relationship questions. The reason for these divergent results is that excluded from motherhood measures derived from household relationship questions is a sizeable sample of biological mothers (and non-biological mothers) who are not co-resident with their children. The excluded sample of not co-resident mothers outweighs the additional number of non-biological mothers or parental surrogates who may be identified.

Table 3. 4: The proportion and counts of all women aged 20 to 49 who are mothers using different identification methods

	PSLSD 1993	OHS 1995	GHS 2002
Method Three – using a maternal relationship question to identify mothers who are co-resident with at least one biological or non-biological child	0.689 (0.008) 5,672,457 (130,571)	0.639 (0.004) 5,450,742 (1,136,141)	0.665 (0.004) 6,647,329 (95,991)
Method Four – using a question on relationships to the household head to identify mothers who are co-resident with at least one biological or non-biological child	0.678 (0.008) 5,581,730 (129,334)	0.660 (0.004) 5,705,415 (1,181,220)	0.642 (0.005) 6,418,282 (93,323)
Method Two – using birth information to identify mothers with at least one biological child who is still alive	0.753 (0.007) 6,202,970 (152,579)	0.706 (0.005) 6,108,159 (1,270,628)	0.762 (0.004) 7,617,630 (102,431)

Notes: Standard errors are in parentheses, counts are on the bottom half of each cell. Estimates are obtained by accounting for weighting and clustering in each survey. The sample includes all women aged 20 to 49. The PSLSD 1993 captured information for both resident and non-resident women but the sample above is limited to resident women only as birth status was only asked of resident household members. No estimates are presented for women in the GHS 2002 with race reported as "other" or "missing".

Table 3. 5: The proportion and counts of women aged 20 to 49 who are mothers using different identification methods, by race

	PSLSD 1993	OHS 1995	GHS 2002
		African	
Method Three – using a maternal relationship question to identify mothers who are co-resident with at least one biological or non-biological child	0.692 (0.008) 4,143,783 (168,167)	0.616 (0.005) 4,010,303 (782,400)	0.659 (0.004) 5,056,660 (90,457)
Method Four – using a question on relationships to the household head to identify mothers who are co-resident with at least one biological or non-biological child	0.675 (0.009) 4,073,403 (165,618)	0.650 (0.004) 4,231,724 (818,467)	0.635 (0.005) 4,876,118 (89,159)
Method Two – using birth information to identify mothers with at least one biological child who is still alive	0.780 (0.006) 4,703,674 (186,289)	0.698 (0.005) 4,545,148 (884,915)	0.770 (0.004) 5,907,002 (99,106)
		Coloured	
Method Three – using a maternal relationship question to identify mothers who are co-resident with at least one biological or non-biological child	0.734 (0.018) 542,829 (168,167)	0.693 (0.009) 585,725 (155,274)	0.752 (0.110) 719,116 (39,783)
Method Four – using a question on relationships to the household head to identify mothers who are co-resident with at least one biological or non-biological child	0.749 (0.017) 553,709 (98,836)	0.706 (0.009) 596,316 (159,058)	0.708 (0.012) 676,405 (35,289)
Method Two – using birth information to identify mothers with at least one biological child who is still alive	0.751 (0.017) 555,253 (100,944)	0.743 (0.008) 627,425 (167,446)	0.806 (0.010) 770,444 (41,234)

Table 3.5 continued	PSLSD 1993	OHS 1995	GHS 2002
		Indian	
Method Three – using a maternal relationship question to identify mothers who are co-resident with at least one biological or non-biological child	0.682 (0.035) 175,865 (52,321)	0.687 (0.019) 175,635 (31,972)	0.686 (0.023) 188,971 (20,405)
Method Four – using a question on relationships to the household head to identify mothers who are co-resident with at least one biological or non-biological child	0.668 (0.029) 172,304 (51,289)	0.683 (0.018) 174,704 (31390)	0.654 (0.027) 179,983 (19,543)
Method Two – using birth information to identify mothers with at least one biological child who is still alive	0.709 (0.037) 182,987 (54,454)	0.721 (0.016) 184,337 (33,078)	0.725 (0.023) 199,545 (21,557)
		White	
Method Three – using a maternal relationship question to identify mothers who are co-resident with at least one biological or non-biological child	0.647 (0.027) 779,980 (113,089)	0.655 (0.012) 679,078 (235,047)	0.628 (0.017) 679,573 (40,252)
Method Four – using a question on relationships to the household head to identify mothers who are co-resident with at least one biological or non-biological child	0.649 (0.027) 782,314 (113,756)	0.678 (0.011) 702,669 (243,086)	0.632 (0.017) 683,204 (40,796)
Method Two – using birth information to identify mothers with at least one biological child who is still alive	0.631 (0.024) 761,056 (110,597)	0.725 (0.011) 751,248 (260,171)	0.681 (0.016) 736,657 (42,006)

Notes: Standard errors are in parentheses, population counts are on the bottom half of each cell. Estimates are obtained by accounting for weighting and clustering in each survey. The sample includes all women aged 20 to 49. The PSLSD 1993 captured information for both resident and non-resident women but the sample above is limited to resident women only as birth status was only asked of resident household members. No estimates are presented for women in the GHS 2002 with race reported as "other" or "missing".

Consistent with the national sample results in Table 3.4, Table 3.5 shows that among African, Coloured and Indian women, motherhood measures are also lower when household relationship questions are used compared to birth status information. Compared

to method two measures that use birth status information, measures derived from method three and four that use household relationship questions are particularly low among Africans. The exception, however, is for White women where motherhood measures in 1993 were higher using the maternal relationship question. A possible reason for this is that method two underestimates the true number of White biological mothers in 1993 due to high levels of missing birth status information in the PSLSD 1993.

In general the results in Table 3.4 and 3.5 provide evidence that excluding not co-resident mothers would significantly underestimate the true number of mothers in South Africa. For this reason the first set of measures which use birth status information to identify biological mothers provide more accurate measures of mothers in South Africa.

However birth status information should also be used in conjunction with household relationship information to distinguish between mothers who are co-resident and not co-resident with their children. In Table 3.6, for example, I measure how many and what proportion of mothers are *not* living with any of their biological children (i.e. are absent mothers) and how this has changed over time. This is achieved by subtracting estimates using method three from estimates using method two (see Tables 3.4 and 3.5). The remainder of women are likely to be biological mothers who are not living with their children i.e. absent mothers. It must be noted that the figures in Table 3.6 represent lower bound estimates of the number of all absent mothers because some non-biological mothers, who are not co-resident with their children, will not be identified.

Table 3.6 shows clear differences in co-residency across race groups. The incidence of mothers not living with their children is highest among the African sample and lowest among Indians and Whites. Therefore a co-residency restriction on motherhood is likely to underestimate the extent of motherhood particularly among Africans. In 2002, 11 percent of all African mothers were not co-resident with their children compared to about five percent of Coloured mothers and White mothers and four percent of Indian mothers. This result is consistent with a well-developed literature on variations in family and household living arrangements across race groups in South Africa (cf. Amoateng *et al* 2007; Zulu and

Sibanda 2005; Russell 2004). These variations in living arrangements are consequences of different socio-economic factors, political histories and cultural ideals, in particular norms regarding co-residence with relatives (Amoateng *et al* 2007; Russel 2004). Magwaza (2003) notes that historically higher incomes among White households compared to African households have afforded White mothers more opportunity to stay at home and care for their children compared to African mothers. Compared to White women, lower marital rates among African women (Kalule-Sabiti *et al* 2007) also result in less financial support from spouses which could be used in raising children. This increases the need for African mothers to migrate to places of employment leaving their children behind in the household of origin.

Table 3. 6: Percentage of women who have one biological child (who is still alive) but are *not* living with any of these biological children

	1993	2002	Percentage point increase from 1993 to 2002
African	8.8 560,936	11.1 850,342	+2.3
Coloured	1.7 12,406	5.4 51,328	+3.7
Indian	2.7 7,122	3.9 10,574	+1.2
White	-	5.3 57,084	-

Source: PSLSD 1993 and GHS 2002.
Notes: The sample is of all women aged 20 to 49 who have had at least one live birth. The PSLSD 1993 sample is limited to resident women only as birth status was only asked of resident household members. Results are not shown for White women in columns two and four due to unreliable estimates of the proportion of women who had experienced a live birth in the PSLSD 1993.

The results in Table 3.6 also suggest that relatively more mothers are living apart from their children in post-apartheid South Africa. The last column shows that among the African, Coloured and Indian samples[15] more mothers were *not* living with their biological children

[15] Trends in living arrangements are not discussed for White women because White biological mothers were considerably underestimated in the presence of missing birth information in the PSLSD 1993.

in 2002 compared to 1993. There was a 2.3 percentage point increase in African mothers (representing an additional 290,000 biological mothers) who were not co-resident with their children from 1993 to 2002. Compared to Africans, the percentage point increase in Coloured women not living with their children was greater at 3.7 percent but this increased from a much smaller absolute base. A possible reason for the increase in the proportion of mothers not living with their children may be higher levels of female migration and labour force participation over this period (Posel and Casale 2006; Casale and Posel 2003). Increasing proportions of mothers may be moving away from home to look for work, leaving children behind. Research by Ardington and Case (2007) and Posel *et al* (2006) suggests that increased female migration may also be related to access to old age pensions within the household. Pension households provide support for women to leave children at home while they migrate to work or to search for work.[16]

Another reason may be falling marital rates among women (Kalule-Sabiti *et al* 2007) and the rise in female-headed households (Posel 2001). Without the financial support of a spouse or male head, women may be pushed into the labour force, leaving their home in search of work. In the face of resource deprivation mothers may be forced to foster out their children or place them in the care of other relatives (McDaniel and Zulu 1996; Van der Waal 1996). The AIDS epidemic could also be another contributing factor. Findings show that mothers who are too ill to care for their children or over-burdened in caring for ill household members may abandon their children or place them in foster care (Anderson and Phillips 2006).

The above analysis highlights that a sizeable sample of mothers, specifically African mothers, are not co-resident with their children. Measures that place a co-residency restriction on motherhood therefore considerably underestimate the extent of motherhood in South Africa, particularly among Africans. Furthermore, the analysis suggests that

[16] In contrast to the results of Ardington and Case (2007) and Posel et al (2006), Maitra and Ray (2003) using the KwaZulu-Natal Income Dynamics Study find that the social pension is negatively related to the number of children in the household. This suggests that the social pension actually decreases the likelihood that children are left with older women (grandmothers) while mothers migrate elsewhere to work.

measures that place a co-residency restriction on motherhood increasingly underestimate motherhood in South Africa because a growing number of mothers are living apart from their children.

3.4 Measuring co-resident mothers: the maternal relationship question versus questions on relationships to the household head

The above discussion has highlighted that using household relationship questions to match women to their children underestimates mothers in South Africa. However the analysis above also showed that if household relationship questions are used in conjunction with birth status information it is possible to distinguish co-resident mothers from absent mothers. In this section I show that if household relationship questions are used to identify co-resident mothers, the maternal relationship question (method three) offers more accurate measures than the question on relationships to the household head (method four).

Compared to the maternal relationship question, questions on relationships to the household head may not adequately identify the motherhood status of women in a household. This is especially the case where mothers live in extended households and are not spouses or parents of the household head (see Chapter Two for more details).

I am able to test whether method four underestimates the number of co-resident biological[17] mothers in the household relative to method three. This is done by comparing measures derived from different combinations of methods. In Table 3.7 I present the proportion of women who are mothers identified by i) combining method two with method three and ii) combining method two with method four. In other words I identify a woman as a mother if she is identified as such by *both* method two *and* method three/four. Both combinations should identify biological mothers who co-reside with at least one living child, but the results are consistently larger for the first combination (i) than the second combination (ii). These results confirm that method four underestimates the number of biological mothers

[17] However this methodology does not allow one to determine if method four underestimates the number of non-biological mothers in the household relative to method three.

relative to method three. Questions on a household member's relationship to the household head cannot be used to accurately match mothers to their biological children living in the same household.[18] Therefore the maternal relationship question provides more reliable measures of co-resident biological mothers than questions on relationships to the household head.

Table 3. 7: Measuring co-resident (biological) mothers – a comparison of method three and four, national sample of women aged 20 to 49

	1993	1995	2002
Combination i) Method two and *three*	0.666 (0.008)	0.625 (0.005)	0.660 (0.004)
Combination ii) Method two and *four*	0.609 (0.008)	0.605 (0.005)	0.591 (0.004)

Source: PSLSD 1993, OHS 1995, GHS 2002.
Notes: Standard errors are in parentheses. Estimates are obtained by accounting for weighting and clustering in each survey. The sample includes all women aged 20 to 49. The PSLSD 1993 sample is limited to resident women only as birth status information was only asked of resident household members.

Conclusion

Measures of motherhood in South Africa vary considerably across different methods used to match women to their children. In this chapter I considered two sets of measures. The first set used birth data to match women to their biological children, regardless of co-residency with children. The second set used questions on household relationships to identify biological and non-biological mothers who are co-resident with at least one of their children. Although the first set of measures does not identify non-biological mothers or women who act as parental surrogates, they consistently yield larger measures of mothers

[18] Posel and Devey (2005:40-41) reach a similar conclusion in their study that counts the numbers of fathers in South Africa. They report much lower percentages of men who are fathers using this method compared to the results of other methods used to match fathers to their children. The reason that questions on relationships to the household head underestimate mothers is that they are used to identify mothers indirectly. Because relationship codes are not sufficiently disaggregated, it's possible that a significant proportion of mothers will not be identified. Also it means that counts of non-biological mothers will also be under-estimated.

than the second set of results. The reason is that among mothers, a significant sample does not live with their children. Using household relationship questions, and specifically questions on relationships to the household head, will underestimate measures of mothers because they can only link mothers to their children if they live together in the same household.

In 2002, 11 percent of African mothers aged 20 to 49 (representing about 850,000 women) were not co-resident with any of their biological children. Among other race groups, the relative proportions and absolute counts of mothers who were not living with children were much smaller: 5.4 percent among Coloured mothers (51,000), 3.9 percent among Indian mothers (10,500) and 5.3 percent among Whites mothers (57,000). Using household relationship questions to measure mothers will therefore particularly underestimate Africans mothers.

Furthermore, using household relationship questions to identify mothers would increasingly underestimate mothers in South Africa given the rise in the proportion of mothers who were not living with their children from 1993 to 2002. For this reason, motherhood measures based on birth status information are more appropriate to identify mothers in South Africa.

However motherhood measures based on birth status information are often too 'noisy' across the different surveys to seem credible. Measures generated from the OHS 1995, the OHS 1997 and the Census 2001, in particular, were inconsistent both with each other and with other surveys. Measures of biological mothers, however, appeared more 'robust' across the PSLSD 1993, the Census 1996 and the GHS 2002. If mothers are identified as women with at least one biological child who is still alive, regardless of co-residency status with children, then about 75 percent of a national sample of women aged 20 to 49 were biological mothers in 1993. This percentage decreased slightly in 1996 to about 72 percent but increased again to 76 percent in 2002.

Chapter Four:

Describing the characteristics of mothers and non-mothers

Introduction

The primary analytical objective of this thesis is to investigate the relationship between motherhood and labour force participation in South Africa. This chapter begins this analysis as I explore descriptively whether mothers differ from non-mothers in terms of their observable characteristics and I identify correlates of labour force participation. It is not possible using descriptive analysis to isolate the effects of motherhood on female labour force participation, while holding all other factors constant. However descriptive statistics provide initial insights into this relationship.

A key research question in the thesis asks whether surrogate parenthood or foster care constrains the labour force participation of women, including non-mothers, in South Africa. Descriptive statistics on labour force participation and other observed characteristics, across non-mothers who do and do not reside with children, provide some initial answers to this research question. The chapter also describes the average characteristics for the sample of mothers and non-mothers which is subsequently used in the econometric analysis in Chapter Five.

Section 4.1 outlines the data and sample used in this chapter (and in Chapter Five) and describes motherhood status among the sample. Section 4.2 compares the average characteristics of non-mothers with a broad category of mothers (not distinguished by co-residency with children), where focus is given to describing differences in their household living arrangements and educational attainment.

Chapter Three suggested that a sizeable sample of all African mothers aged 20 to 49 did not live with their children in 2002. These mothers are likely to be labour migrants who have left their children in the household of origin. They may also differ significantly in terms of their observed (and unobserved) characteristics compared to mothers who are co-resident with their children. In section 4.3 I describe and compare the measured characteristics of these two groups of mothers. Attention is given particularly to examining differences in the labour force activity of non-mothers and mothers, distinguished by co-residency with children. Section 4.4 then considers how these differences map onto differences in other observed characteristics of these non-mothers and mothers. Finally, I compare average characteristics, specifically labour force participation, across non-mothers living with and without children in section 4.5. I examine whether there is descriptive evidence that co-residence with children (and the provision of childcare) constrains the labour force participation of *non-mothers*.

4.1 Data and definitions

In this thesis I use the nationally representative General Household Survey (GHS) 2002 to investigate the relationship between motherhood and labour force participation. The GHS 2002 is the most recent survey which simultaneously captures information on fertility and a range of socio-economic variables. I have limited the sample of analysis to *African* women only and consistent with the analysis in Chapter Three, I restrict the age of women to 20 to 49 years. I also restrict the definition of motherhood to women with at least one *biological* child aged *18 years or younger*, where the *child is still alive*. A non-mother is a woman with no biological children aged 18 years or younger. The sample restrictions and motherhood definition used in this chapter (and subsequent chapters) are motivated in light of the specific research questions and the overall objective of this thesis.

The primary objective of this thesis is to explore the economic costs of motherhood/childbearing in South Africa; and specifically if decisions to actively participate in the labour market are constrained by childbearing or childcare responsibilities. A specific research question asks how restricting the definition of

motherhood to only co-resident mothers biases the estimated relationship between motherhood and women's labour force participation. A second research question asks if living with children who are not one's own and surrogate parenthood also impacts on women's labour force participation.

The first research question particularly applies to an African sample of women rather than all women. Among all women in South Africa, the effect on labour force participation estimations of restricting the definition of motherhood to co-resident mothers only is likely to be largest for African women. The reason is that they are more likely than women in other race groups to be living apart from their children. This was explored in detail in Chapter Three which showed that the incidence of mothers not co-residing with their children was low among Coloured, Indian and White women but high among African women (see also Appendix 4A). Furthermore, African women in South Africa are more likely than all other race groups to be living with children who are not their own in extended family structures (see Appendix 4A). Thus the second research question, which asks whether women's labour force participation is constrained by caring for children who are not one's own, is also chiefly applicable to an African sample of women. The restriction on African women also simplifies the investigation of the effects of childbearing on women's labour force participation because mothering and childcare practices vary widely across different race groups (Russell 2004; Magwaza 2003).

In the previous two chapters, I did not place a restriction on the age of children belonging to a mother. The focus in those chapters was to explore motherhood from a population studies or demographic perspective. In this chapter however, the focus shifts to exploring the economic costs of motherhood/childbearing which may be sensitive to the age of a mother's children. Adult children may not impose either a pecuniary or indirect economic cost upon their parents. In some cases, adult children may in fact be an economic benefit rather than a cost to their parents as they provide financial support. Accordingly, motherhood is restricted to women with children 18 years or younger. Women with children older than 18 years are not included in the sample of mothers.

I also restrict mothers to women with *biological* children. There are two reasons for this. First, I am unable to identify non-biological mothers who are not co-resident with their children in nationally representative household surveys. By restricting the sample to biological mothers only I therefore generate a consistent sample of mothers, regardless of co-residency with children. Second, the international literature about the effects of motherhood on women's labour force participation, typically only identifies biological mothers. For the purpose of comparing my results with international findings, I also use this restriction.

In the chapter I distinguish between mothers who do and do not co-resident with their children aged 18 years or younger. A co-resident mother is a woman living with at least one biological child 18 years or younger in the same household. A not co-resident mother is a woman with at least one biological child 18 years or younger but who does not live with any of her children. This distinction is made using a question in the birth module which asked the mother if she lived with her children in the household. This information is then corroborated using a maternal relationship question in the household roster which asked all household members to identify his/her mother if she lived in the same household. In principle, the birth module information should be sufficient to identify co-residency. However, 83 observations reported living with biological children in the household while no one in the household reported these women to be their mothers. One explanation for this is that some mothers accidentally reported living with their children. However coding errors are most likely to be the reason for the discrepancy[19]. To eliminate inconsistencies in data reporting, a mother is defined as co-resident with at least one child if she reports living with the child *and* at least one household member reports that she is his/her mother. Therefore the 83 discrepant observations are not included in the sample of mothers.

[19] Another less likely explanation is that these children are raised by another individual within the same household who is not their biological mother. For example, in the case of adolescent pregnancies the adolescent's mother may become the primary caregiver to her daughter's child.

Throughout this chapter and subsequent chapters, I also exclude from the category of mothers those women whose every child has died after birth.[20] Without this exclusion two types of not co-resident mothers would be aggregated together: 'absent' mothers and mothers whose children have died.[21] Any conclusions on the characteristics of 'absent' mothers will be confounded if they are not distinguished from the latter.

Table 4.1: Motherhood and co-residency with children, African women and children 2002

	Among all women				Among co-resident mothers	Among all children
	Mother, co-resident with own children	Mother, not co-resident with own children	Not mother, co-resident with other children	Not mother, not co-resident with other children	Co-resident with own and not own children	Not co-resident with mother
Proportion	0.613 (0.005)	0.108 (0.003)	0.191 (0.004)	0.089 (0.003)	0.463 (0.007)	0.279 (0.005)
N	10,015	1,998	3,348	1,674	4,634	10,085
Rural						
Proportion	0.654 (0.007)	0.097 (0.005)	0.189 (0.005)	0.060 (0.003)	0.519 (0.010)	0.313 (0.006)
N	4,963	862	1,565	569	2,538	6,636
Urban						
Proportion	0.573 (0.007)	0.118 (0.005)	0.191 (0.005)	0.117 (0.005)	0.403 (0.011)	0.224 (0.006)
N	5,052	1,136	1,783	1,105	2,096	3,449

Source: GHS 2002, Posel and van der Stoep 2008.
Notes: The data are weighted and adjust for clustering and stratification. Sample sizes are not weighted. Standard errors are in parentheses. The data for mothers and non-mothers are for women aged 20 to 49 and includes those women currently in school. Children are aged 18 years or younger.

[20] Among women who had reported a live birth, 123 women reported that all their children had subsequently died since birth. These women were not classified as mothers.

[21] If mothers whose children have died since birth are not excluded from the sample of mothers they will be identified as mothers whose children do not live in the same household.

In Table 4.1, I describe motherhood status among the sample of African women aged 20 to 49 used in this study. In the first four columns I distinguish mothers and non-mothers according to co-residency with children. Among African women aged 20 to 49, 61 percent are co-resident mothers (10,015) and 11 percent are not co-resident mothers (1,998) in 2002. A further 19 percent are non-mothers but are living in the same household with children who are not their own. The remaining nine percent are non-mothers who are not living with any children. Among co-resident mothers specifically, about 46 percent are living also with other children who are not their own (Posel and van der Stoep 2008). The high incidence of co-residency with other children among the African sample is consistent with extended family structures and child fosterage in African households (Amoateng *et al* 2007; Anderson and Phillips 2006; Russell 2004; Van der Waal 1996).

The results in Table 4.1 also reflect a rural/urban dimension to living arrangements and motherhood status in South Africa. A large proportion of all African women in rural areas are mothers who are co-resident with their children: 65 percent in rural areas compared to 57 percent in urban areas (see data column one). This finding mirrors research on fertility rates in South Africa which are higher in rural compared to urban areas (Moultrie and Timaues 2001; Mencarini 1999; Mostert and Hofmeyer 1988). African women in rural areas are also significantly more likely to be co-resident with other children. For example 52 percent of co-resident mothers in rural areas, compared to 40 percent in urban areas, were also living with children who were not their own (see data column five). These results are consistent with the co-residence of extended families in rural households in particular (Posel and van der Stoep 2008; Amoateng *et al* 2007:54).

Not co-resident mothers are more likely to be living in urban areas (56 percent or 1,136/1,998). Analogously, a larger a proportion of children in rural areas report not living with their mother. This is consistent with patterns of labour migration to places of employment, particularly from rural to urban areas (Posel and van der Stoep 2008; Posel and Casale 2006). It may also reflect a traditional practice among African urban residents to send their children to their parents and grandparents in rural areas for care (Amoateng *et al* 2007:52).

In analysing the relationship between motherhood/childbearing and labour force participation, research conclusions may be particularly sensitive to the sample of women selected. In the above analysis which described motherhood status among African women, I included those women who are currently in school; but in the following analysis I exclude those women who are not currently in school[22] to retain a more comparable sample of mothers and non-mothers. About 15 percent of African non-mothers aged 20 to 49 were still in school in 2002 compared to two percent of mothers. Individuals who are still in school do not constitute *potential* labour force participants (Casale 2003:119). If included in the sample they may downwardly bias the labour force participation rates of non-mothers and bias the estimation of the motherhood effect in Chapter Five. Given this restriction, I identify 4,061 non-mothers and 11,364 mothers (including 9,488 co-resident mothers and 1,876 not co-resident mothers).

4.2 Comparing the characteristics of mothers and non-mothers

In this section I analyse differences in observed characteristics across non-mothers and mothers, regardless of their co-residency status with children. I first describe and compare their demographic characteristics and household living arrangements, followed by their educational attainment.

4.2.1 Demographic characteristics and household living arrangements

Table 4.2 presents the average characteristics of non-mothers and mothers (women with at least one biological child 18 years or younger, regardless of co-residency with children). The sample of mothers here combines co-resident and not co-resident mothers into a single category.

[22] Excluding the currently in school results in an older sample of non-mothers who are less likely to be living with parents, and more likely to be married and to have slightly less education.

The table shows that mothers are significantly older than non-mothers. This is consistent with international studies on motherhood (Boushey 2008; Budig and England 2001). The age distribution of non-mothers is skewed toward the youngest age cohorts suggesting that the majority of these women may be women who have delayed entry into motherhood rather than deciding not to have children at all. The average age of entry into motherhood (i.e. age at first birth[23]) is 21 years for this sample of African women.

Literature on the determinants of fertility predicts a positive relationship between marriage and children ever born. South African studies specifically find that mothers are more likely to be married compared to childless women (Moultrie and Timaues 2001:217). Accordingly, the results in Table 4.2 show that marital rates are significantly higher among mothers than non-mothers. In 2002, about fifty percent of mothers were married or living with a partner compared to 21 percent of non-mothers. Analogously, a significantly smaller proportion of mothers had never been married (42 percent compared to 69 percent of non-mothers).

Moultrie and Timaues (2001) explore the living arrangements of women in South Africa. Using the 1993 Project for Statistics on Living Standards and Development (PSLSD), they find that women's living arrangements are contingent on the stage of the life-course that they are experiencing and differ by her motherhood status (*ibid*:215). In particular mothers are less likely to be living with their parents or people of their own generation even when partially controlling for age (*ibid*:217). Results presented in Table 4.2 using the GHS 2002 agree with the findings of Moultrie and Timaues (2001). On average mothers are significantly less likely to be living with their parents (grandparents of their children). Mothers also live with fewer numbers of economically inactive females and fewer pension-aged men and women. Despite these results, mothers on average live in larger households due to the presence of own children aged zero to 18 years and the increased likelihood of living with a spouse. In contrast, non-mothers on average live with more children aged zero to 18 who are *not* their own children.

[23] The GHS 2002 birth module did not directly ask a question on age at first birth, however it was determined by subtracting the age of a mother's first child from her own reported age.

Table 4. 2: Demographic and household characteristics of African non-mothers and *all* mothers, 2002

	Non-mothers		All Mothers	
Rural	0.432	(0.012)	0.510*	(0.009)
Age at first birth		-	21.077*	(0.054)
Age	30.621	(0.167)	33.004*	(0.090)
Age distribution:				
20 – 24	0.356	(0.010)	0.157*	(0.004)
25 – 29	0.220	(0.008)	0.226	(0.005)
30 – 34	0.108	(0.006)	0.209*	(0.005)
35 – 39	0.079	(0.005)	0.165*	(0.004)
40 – 44	0.103	(0.006)	0.147*	(0.004)
45 – 49	0.133	(0.006)	0.096*	(0.003)
Marital Status:				
Married and co-resident with spouse	0.204	(0.008)	0.391*	(0.007)
Married but not co-resident with spouse	0.044	(0.004)	0.115*	(0.004)
Divorced / separated / widowed	0.061	(0.004)	0.075	(0.003)
Never Married	0.691	(0.009)	0.420*	(0.007)
Household composition and resources:				
Household size	5.387	(0.081)	5.967*	(0.058)
Women's mother (grandmother to women's child) in household	0.370	(0.010)	0.239*	(0.006)
Women's father (grandfather) in household	0.173	(0.009)	0.098*	(0.004)
Co-residing with at least one biological child		-	0.851	(0.005)
Connected to piped water	0.605	(0.013)	0.542*	(0.010)
Connected to electricity	0.732	(0.012)	0.707	(0.010)
Total monthly earnings per adult equivalent	477.099	(21.741)	420.521	(17.162)
Number co-resident in household:				
Economically inactive women over 18	0.521	(0.019)	0.363*	(0.012)
Economically inactive men over 18	0.390	(0.016)	0.320*	(0.010)
Pension-aged men	0.081	(0.006)	0.064	(0.003)
Pension-aged women	0.195	(0.009)	0.181	(0.006)
Employed adult men	0.426	(0.016)	0.445	(0.010)
Own children 0 to 18		-	1.816	(0.017)
Other children 0 to 18	1.838	(0.047)	1.147*	(0.030)
N	4,061		11,364	

Source: GHS 2002

Notes: Mothers are women who have at least one biological child aged 18 years or younger, regardless of co-residency with children. The data are weighted and control for clustering and stratification in sample design. Standard errors are in parentheses. The sample includes African women aged 20 to 49 and excludes those women currently in school. Earnings are in 2002 prices and do not include non-earned income. "Total monthly earnings" are normalized by the adult equivalent household size which is constructed by signing a value of one to individuals over the age of 11 while children under the age of 12 are assigned a value of 0.5.

* indicates a significant difference in means and proportions across non-mothers and all mothers using a 95 percent confidence interval.

Table 4.2 also illustrates differences in motherhood by location: in 2002, 51 percent of mothers lived in rural areas compared to 43 percent of non-mothers. This is in line with both international and South African literature that finds higher fertility rates among women living in rural areas (Moultrie and Timaues 2001; Mencarini 1999). Location, in turn, is highly correlated with access to services and resources. Compared to non-mothers, mothers are significantly less likely to be living in households with access to piped water. Mothers also live in households with lower average monthly earned income per adult equivalent; however the difference in incomes is not statistically significant at the 5 percent level of significance.

4.2.2 Educational attainment

Both the theoretical and empirical literature on motherhood suggests that mothers possess less education than non-mothers (Piras and Ripani 2005; Anderson *et al* 2003; Budig and England 2001; Marini 1984; Mincer and Polachek 1974). Marini (1984) argues that mothers may have lower educational attainment compared to non-mothers "either because an unplanned birth forces premature termination of schooling or because a planned birth at a relatively young age forecloses the option of returning to school after leaving (Marini 1984:493)". Alternatively, Mincer and Polachek (1974) argue that lower education among mothers can be explained in terms of returns to education. Women who desire to have a child may be inclined to invest in less education as the length of the expected payoff period on their investment is shorter. Years spent out of the labour market while rearing children reduce the period in which to acquire higher earnings associated with investments in higher education.

Table 4.3 compares educational attainment across mothers and non-mothers in South Africa in 2002. The data show that mothers have 0.4 years less of schooling on average than non-mothers. In particular, mothers are significantly less likely to have completed

secondary schooling (or matric) compared to non-mothers.[24] However it is not possible here to determine if lower educational attainment among mothers is the cause or effect of childbearing. Entry into motherhood may constrain further educational attainment. Alternatively, women with lower levels of education may be more likely to enter motherhood because they forgo lower returns to education in the market place compared to women with higher levels of education.

Table 4. 3: Educational attainment among non-mothers and *all* mothers, African women 2002

	Non-mothers	All mothers
Completed years of schooling	8.806 (0.085)	8.397* (0.056)
No schooling	0.089 (0.006)	0.083 (0.003)
Primary	0.222 (0.008)	0.267* (0.006)
Incomplete secondary	0.335 (0.009)	0.399* (0.006)
Matric (completed secondary)	0.279 (0.010)	0.183* (0.005)
Postmatric	0.075 (0.006)	0.068 (0.004)
N	4,061	11,364

Source: GHS 2002

Notes: Mothers are women who have at least one biological child 18 years or younger, regardless of co-residency with children. Primary schooling includes schooling up to grade 7; incomplete secondary includes schooling from grade 8 to 11; matric represents completed grade 12 education. The data are weighted and control for clustering and stratification in sample design. Standard errors are in parentheses. Sample sizes are not weighted. The sample includes African women aged 20 to 49 but excludes those women currently in school. * indicates a significant difference in mean years of schooling and the proportion with a specific level of schooling attainment across non-mothers and all mothers using a 95 percent confidence interval.

South African literature has documented rising levels of education among women in South Africa (Casale 2003:102; Mlantsheni and Leibbrandt 2001). In Table 4.4 I describe how educational attainment has changed by age cohort of women. The table presents average years of completed schooling for non-mothers and mothers across age cohorts. Differencing average years of schooling for women in age cohort 45 to 49 from women in

[24] Mothers are also less likely to be currently enrolled in school: if all women aged 20 to 49 are included in the sample then only five percent of mothers can be identified as being enrolled in an educational institution compared to 18 percent of non-mothers (results not presented).

age cohort 25 to 29,[25] highlights by how much educational attainment has increased in both samples of mothers and non-mothers. However the positive education differential across young and old cohorts is larger among non-mothers than mothers. Educational attainment has increased by 4.6 years among non-mothers compared to only 3.5 years among mothers. Comparing average years of schooling across mothers and non-mothers, there are no significant differences at older age cohorts (34 to 49), but at age cohort 20 to 29 years of schooling are significantly higher among non-mothers. Thus educational attainment of mothers and non-mothers has diverged in recent years. With greater access to education in post-apartheid South Africa, especially for African women, the opportunity cost of motherhood in terms of forgone education appears to be rising.

Table 4. 4: Average years of schooling, African women 2002

	Non-mothers	All mothers
20 – 24	10.174 (0.100)	9.298* (0.086)
25 – 29 (1)	10.240 (0.130)	9.515* (0.094)
30 – 34	8.812 (0.254)	9.102 (0.106)
35 – 39	7.076 (0.257)	7.958 (0.117)
40 – 44	6.331 (0.236)	6.779* (0.124)
45 – 49 (2)	5.656 (0.212)	5.992 (0.144)
All women 20 – 49	**5.992 (5.992)**	**8.397* (0.056)**
Mean difference (1) – (2)	4.584	3.523
N	4,061	11,364

Source: GHS 2002

Notes: Mothers include all women who have at least one biological child aged 18 years or younger, regardless of co-residency with children. The data are weighted and control for clustering and stratification in sample design. Standard errors are in parentheses. The sample includes African women aged 20 to 49 and excludes those women currently in school. * indicates that mean years of schooling are significantly different across non-mothers and mothers using a 95 percent confidence interval.

[25] The educational attainment of women in age cohort 20 to 24 is less than women aged 24 to 29. In part this reflects that women in the younger age cohort are more likely to include women who have not yet completed their education. Furthermore, lower levels of educational attainment among *mothers* aged 20 to 24 compared to age 25 to 29 could be attributed a negative causal relationship between age at first birth and education. The younger age cohort captures a larger portion of mothers who have a child at an early age who are more likely to truncate their education (Marini 1984:492).

4.3 Comparing the characteristics of non-mothers, co-resident and absent mothers

In this section, I now differentiate among mothers according to whether or not they co-reside with at least one of their children aged 18 years or younger.

4.3.1 Demographic characteristics and household living arrangements

Table 4.5 presents the average characteristics of all non-mothers and the two samples of mothers: co-resident mothers and not co-resident mothers. The samples co-respond to the samples used in the econometric analysis of the next chapter (Chapter Five).

The table presents distinct differences in the average individual and household characteristics of co-resident and not co-resident mothers. Co-resident mothers are on average significantly older than not co-resident mothers (33.2 years compared to 31.7 years) with an older age at first birth. Compared to not co-resident mothers, co-resident mothers on average also have more biological children and are more likely to have younger children, specifically under the age of four. Co-resident mothers are also more likely to be married than not co-resident mothers: in 2002 about sixty percent of co-resident mothers were married or had ever been married compared to about 45 percent of not co-resident mothers. The majority of not co-resident mothers are women who have never been married (55 percent). However, on average there are no significant differences in the educational attainment of these two groups of mothers using a 95 percent confidence interval.

The table also shows that there are differences in location and variations in household living arrangements across co-resident and not co-resident mothers. Co-resident mothers are significantly more likely to live in rural areas and in larger households with more adults, pensioners and children who are *not* their own. They are therefore more likely to be living in extended households compared to not co-resident mothers.

The above discussion has highlighted many significant average differences in characteristics across co-resident and not co-resident mothers. This suggests that the allocation of mothers across these two groups is not random.

Table 4. 5: Demographic and household characteristics of non-mothers, co-resident mothers and not co-resident mothers, African women 2002

	Non-mothers		Mothers: co-residing with at least one (own) child		Mothers: not co-resident with (own) children	
Rural	0.432	(0.012)	0.523	(0.009)	0.433*	(0.017)
Age	30.621	(0.167)	33.224	(0.099)	31.745*	(0.204)
Age at first birth	-		21.207	(0.058)	20.336*	(0.123)
Marital Status:						
Married and co-resident with spouse	0.204	(0.008)	0.403	(0.007)	0.323*	(0.014)
Married but not co-resident with spouse	0.044	(0.004)	0.123	(0.005)	0.070*	(0.007)
Divorced / separated / widowed	0.061	(0.004)	0.077	(0.003)	0.060*	(0.006)
Never Married	0.691	(0.009)	0.398	(0.007)	0.546*	(0.015)
Educational attainment:						
Years of schooling	8.806	(0.085)	8.363	(0.061)	8.588	(0.110)
No schooling	0.089	(0.006)	0.085	(0.004)	0.072	(0.007)
Primary	0.222	(0.008)	0.270	(0.006)	0.250	(0.013)
Incomplete secondary	0.335	(0.009)	0.395	(0.007)	0.420	(0.013)
Matric	0.279	(0.010)	0.181	(0.005)	0.198	(0.011)
Postmatric	0.075	(0.006)	0.069	(0.004)	0.059	(0.006)
Household composition and resources:						
Household size	5.387	(0.081)	6.331	(0.059)	3.878*	(0.107)
Women's mother (grandmother to women's child) in household	0.370	(0.010)	0.256	(0.006)	0.140*	(0.010)
Women's father (grandfather) in household	0.173	(0.009)	0.103	(0.004)	0.068*	(0.007)
Connected to piped water	0.605	(0.013)	0.535	(0.010)	0.581	(0.017)
Connected to electricity	0.732	(0.012)	0.713	(0.010)	0.673	(0.018)
Number co-resident in household:						
Economically inactive women over 18	0.521	(0.019)	0.381	(0.012)	0.259*	(0.021)
Economically inactive men over 18	0.390	(0.016)	0.338	(0.010)	0.221*	(0.019)
Employed men over 18	0.426	(0.016)	0.450	(0.011)	0.413	(0.019)
Pension-aged men	0.081	(0.006)	0.066	(0.004)	0.048	(0.006)
Pension-aged women	0.195	(0.009)	0.192	(0.007)	0.113*	(0.010)
Own children 0 to 18 (resident)	-		2.088	(0.017)	-	
Other children 0 to 18 (resident)	1.838	(0.047)	1.176	(0.032)	0.981*	(0.054)
Number of own children 0 to 18	-		2.250	(0.018)	1.172*	(0.029)
Age of youngest child ever born	-		4.925	(0.053)	7.116*	(0.146)
Child under four years	-		0.471	(0.007)	0.275*	(0.013)
N	4,061		9,488		1,876	

Source: GHS 2002

Notes: The data are weighted and control for clustering and stratification in sample design. Sample sizes are not weighted. Standard errors are in parentheses. The sample includes African women aged 20 to 49 who are not currently in school. Children are aged 18 years or younger. * indicates that means and proportions are significantly different across co-resident and not co-resident mothers using a 95 percent confidence interval.

4.3.2 Labour force participation

In developed countries empirical evidence generally finds a negative relationship between motherhood and labour force participation; and labour force participation rates are consistently lower among mothers than non-mothers. This was discussed in Chapter One. Developing country literature, however, provides ambiguous evidence of the relationship between motherhood and labour force participation. Labour force participation has at times been found to be higher among mothers than non-mothers (Bianchi 2000; Piras and Ripani 2005). In the next chapter I explore the relationship between motherhood and labour force participation using econometric techniques. Here I investigate this relationship descriptively.

Throughout the thesis, I use the *strict* definition of labour force participation which defines labour force participants as either employed, or as unemployed individuals who have taken active steps to search for work in the past month. The reason is that I want to investigate whether motherhood constrains women's active participation in the labour market. The non-searching unemployed are therefore included in the category of inactive participants (or not economically active).

Table 4.6 presents labour force participation rates for non-mothers, co-resident mothers and not co-resident mothers. There are larger differences in labour force participation rates *among* mothers than between mothers and non-mothers: in 2002, 77 percent of mothers not living with their children were labour force participants compared to 61 percent of co-resident mothers and 65 percent of non-mothers. Therefore labour force participation rates are highest among not co-resident mothers followed by non-mothers and lowest among co-resident mothers. This result holds at every age cohort. This is seen in Figure 4.1 which plots labour force participation profiles for non-mothers and mothers.

Table 4. 6: Labour force participation rates, African women 2002

	Non-mothers	Mothers: co-residing with at least one (own) child	Mothers: not co-resident with (own) children
Labour force participant (strict)	0.649 (0.010)	0.605 (0.007)	0.765* (0.012)
N	4,061	9,488	1,876

Source: GHS 2002

Notes: The data are weighted and control for clustering and stratification in sample design. Sample sizes are unweighted. Standard errors are in parentheses. The sample includes African women aged 20 to 49 who are not currently in school. Children are 18 years or younger. * indicates that labour force participation rates among not co-resident mothers are significantly higher than those of co-resident mothers using a 95 percent confidence interval.

Figure 4.1 shows a significant divergence in the labour force participation profiles of non-mothers and co-resident mothers at younger age cohorts; but at older cohorts their labour force participation profiles converge. This reflects that as children get older it becomes easier for co-resident mothers to leave their children during the day to work or to search for work. This is confirmed by higher labour force participation rates among co-resident mothers with older children, shown in Table 4.7. The labour force participation rates of co-resident mothers increase particularly when children reach crèche age (four to six years). Among co-resident mothers with children under four years, 53 percent were labour force participants compared to 65 percent of co-resident mothers with children aged four to six. The results of Figure 4.1 and Table 4.7 together suggest that the negative effects of motherhood on labour force participation are driven by the presence of very young children. This is consistent with international literature which finds that labour force participation is increasing in the age of the youngest child (Browning 1992:1450).

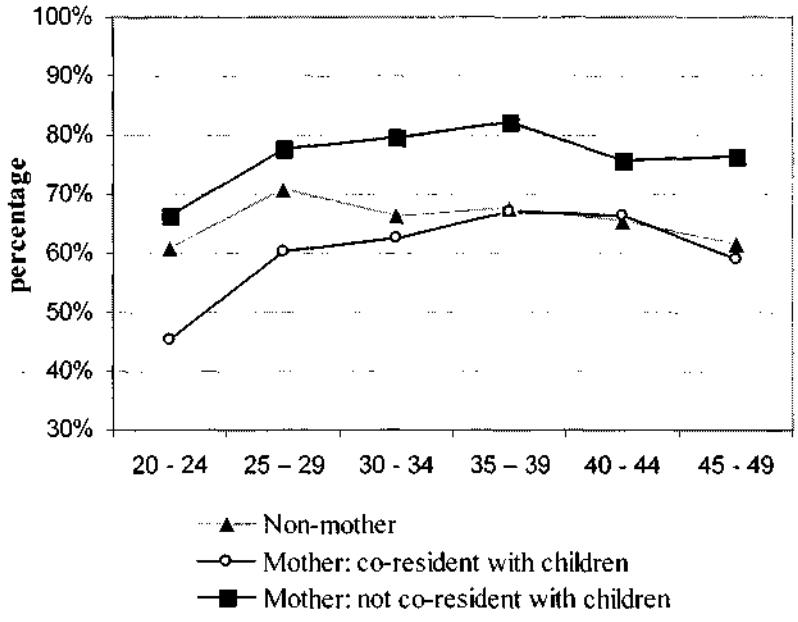
Table 4. 7: Labour force participation rates among African co-resident mothers by age of their youngest child, 2002

0 – 3	0.529 (0.011)
4 – 6	0.648 (0.014)
7 – 12	0.691 (0.011)
13 – 18	0.678 (0.018)

Source: GHS 2002

Notes: The data are weighted and control for clustering and stratification in sample design. Standard errors are in parentheses. The sample includes African co-resident mothers aged 20 to 49 who are not currently in school.

Figure 4. 1: Strict labour force participation profiles by age cohort, African women 2002



Source: GHS 2002 and calculations in Appendix 4B.

4.4 Describing differences in labour force participation rates among mothers

An obvious reason for higher rates of labour force participation among not co-resident mothers, compared to co-resident mothers, is that they do not face direct childrearing responsibilities. This gives them more opportunity to pursue labour market activity. But other differences in their individual characteristics (such as education, marriage and fertility) and household factors could suggest possible explanations for the divergence in their labour force participation. In this section, I describe the correlates of labour force participation and I compare differences across co-resident and not co-resident mothers.[26]

Education is one of the most important determinants of women's labour force participation in both international studies and South African studies (Del Boca and Locatelli 2006;

[26] A complex set of causal relationships likely exists across various socio-economic variables. Therefore causal relationships cannot be inferred from this descriptive analysis.

Casale 2003; Bhorat and Leibbrandt 2001; Winter 1999; Mincer and Polachek 1974). The results in Table 4.8 indicate that labour force participation rates are increasing in education for both co-resident and not co-resident mothers; and among co-resident mothers specifically, the increase in labour force participation rates with education is non-linear. Furthermore the effect of education on labour force participation rates is less dramatic among not co-resident mothers. The results also suggest that at every level of educational attainment up to matric, labour force participation rates are significantly higher among not co-resident mothers than among co-resident mothers.

Table 4. 8: Labour force participation rates among African mothers by their education, 2002

	Mothers: co-residing with at least one (own) child	Mothers: not co-resident with (own) children
No schooling	0.439 (0.022)	0.615* (0.046)
Primary	0.537 (0.013)	0.774* (0.024)
Incomplete secondary	0.594 (0.011)	0.734* (0.020)
Matric (completed secondary)	0.701 (0.014)	0.833* (0.023)
Postmatric	0.873 (0.016)	0.886 (0.040)
N	4,061	9,488

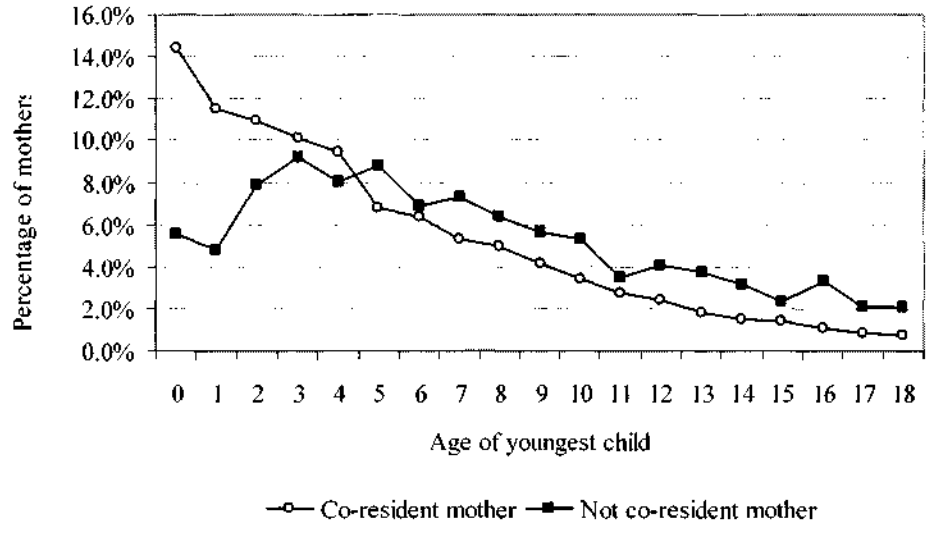
Source: GHS 2002

Notes: The data are weighted and control for clustering and stratification in sample design. Standard errors are in parentheses. Sample sizes are unweighted. Primary schooling includes schooling up to grade 7; incomplete secondary includes schooling from grade 8 to 11; matric represents completed grade 12 education. The sample includes African women aged 20 to 49 with children aged 18 years or younger. Women who are currently in school are excluded. * Indicates that labour force participation rates among not co-resident mothers are significantly higher than those of co-resident mothers using a 95 percent confidence interval.

Fertility and the age of a woman's children are expected to be highly correlated with her labour force participation (Browning 1992). Figure 4.2 shows what proportion of mothers have a youngest child of a specific age under 19 years. The graph clearly illustrates that co-resident mothers are more likely to have a child younger than four years compared with not co-resident mothers. A consistently larger percentage of not co-resident mothers have a child older than four years compared with co-resident mothers. Table 4.9 also indicates that the age of a co-resident mother's youngest child is on average two years greater than for the youngest child of a not co-resident mother. It may be argued that not co-resident mothers

may be more likely to migrate to work because it easier to leave their (on average) older children behind in their household of origin. This argument is supported by Townsend *et al* (2002:214) who find that the likelihood that a child's mothers will be a migrant increases with the child's age using Agincourt data.

Figure 4. 2: The percentage of mothers whose youngest child is of a specific age, African mothers aged 20 to 49, 2002



Source: GHS 2002 and calculations in Appendix 4C.

Table 4. 9: Average age of youngest child and number of live births aged 0 to 18, African mothers, 2002

	Mothers: co-residing with at lest one (own) child	Mothers: not co-resident with (own) children
Age of youngest child ever born	4.925 (0.053)	7.116* (0.146)
Number of live births	2.759 (0.024)	2.112* (0.041)
N	9,488	1,876

Source: GHS 2002

Notes: The data are weighted and control for clustering and stratification in sample design. Sample sizes are not weighted. Standard errors are in parentheses. The sample includes African women aged 20 to 49 who are not currently in school. Children are aged 18 years or younger. * indicates that mean characteristics are significantly different across co-resident and not co-resident mothers using a 95 percent confidence interval.

The descriptive results in Table 4.9 also show that fewer numbers of children are born to not co-resident mothers. Having fewer children would also facilitate migration and entry into the labour market; but caution must be taken when making this inference. Lower fertility among not co-resident mothers may be both a cause and an effect of migration and labour force participation (Jensen and Ahlburg 2004). Fertility differences among mothers may not exist prior to migration (Brockerhoff and Yang 1994 cited in Jensen and Ahlburg 2004);[27] rather increased access to employment at the migrant's destination may increase the opportunity cost of childbearing, lowering her fertility *after* migration (Jensen and Ahlburg 2004:1).

Literature on labour supply also identifies that household factors can have a significant impact on female labour force participation (Casale 2003; Becker 1965). Studies of female labour force participation in South Africa, for example, suggest that reduced access to traditional forms of income support within the household may have '*pushed*' women into the labour market (Casale and Posel 2002; Casale 2003:121). Unfortunately not co-resident mothers, who are labour migrants, cannot be linked back to their household of origin because the GHS 2002 does not collect information on non-resident household members. It is not possible to determine how access to both earned and unearned income in their original households may have influenced their decision to leave home and to find work. However lower marital rates among not co-resident mothers, compared with co-resident mothers, suggest that these women are less likely to be supported by a spouse or partner (see Table 4.5). Not co-resident mothers may have a greater need to work which may require migration to places of employment. Again this argument is also confounded by causality issues. Mothers who have migrated away from their children to work may then have less need to get married because they become economically independent (Casale 2003:120; Gustafson and Worku 2006:21).[28]

---

[27] Brockerhoff and Yang (1994) examine the effects of rural-urban migration on fertility in Sub-Saharan African countries and find *higher* levels of fertility among migrants from rural areas *prior* to migration.

[28] There is evidence of potential endogeneity between marital status and labour force participation in South Africa (Gustafsson and Worku 2006:18). It is expected that higher labour force participation among unmarried mothers may discourage marriage as they are economically independent. This is confirmed by Gustafsson and Worku (2006) who run an ordered probit model for marital choice among South African

This analysis has shown that fewer and older children, as well as lower marital rates, are correlated with higher labour force participation rates among not co-resident mothers, compared to co-resident mothers. Although not observed, household factors in the household of origin are also likely to be correlated with labour force participation among not co-resident mothers. The two groups of mothers may also differ further in terms of their unobserved individual characteristics which may be correlated with their decisions to participate in the labour force. In particular mothers who have moved away from their children may have greater motivation to work, more ambition or a greater need to participate in the labour force than co-resident mothers.

4.5 The characteristics of non-mothers living with and without other children

Given that a sizeable sample of African mothers in South Africa is not co-resident with their children, childcare is probably provided by another relative or household member residing in the household of origin. This is consistent with high levels of fosterage in African households (Van der Waal 1996). Therefore the costs of motherhood may not be borne by mothers alone. In this section I investigate whether childcare costs are borne by non-mothers in particular. I describe and compare the characteristics of two groups of non-mothers, distinguished by whether or not they live with children who are not their own.

A significant proportion of African non-mothers aged 20 to 49 live in households with other children. Of the 4,061 non-mothers identified in section 4.2, 64 percent (2,613/4,061) were co-resident with other children aged 18 years or younger. Among these co-resident non-mothers, specifically 46 percent were living in a household containing at least one child with an absent mother.

---

mothers using the Census 2001. Including an exogenous variable for whether or not a woman is employed, they find that being employed increases the likelihood that mothers will remain single (Gustaffsson and Worku 2006:21).

Table 4. 10: Characteristics of non-mothers who are co-resident and not co-resident with children, African women 2002

	Non-mothers co-resident with other children	Non-mothers not co-resident with other children
Age	29.520 (0.204)	32.793* (0.290)
Daughter/granddaughter of head	0.560 (0.013)	0.164* (0.013)
Rural	0.479 (0.015)	0.343* (0.018)
Marital Status:		
Co-resident with spouse/partner	0.123 (0.008)	0.360* (0.017)
Married but not co-resident with spouse/partner	0.044 (0.005)	0.043 (0.007)
Divorced / separated / widowed	0.057 (0.005)	0.069 (0.008)
Never Married	0.775 (0.010)	0.527* (0.018)
Educational attainment:		
Years of schooling	8.956 (0.102)	8.462 (0.145)
No schooling	0.087 (0.007)	0.094 (0.009)
Primary	0.210 (0.009)	0.244 (0.014)
Incomplete secondary	0.329 (0.011)	0.346 (0.016)
Matric	0.300 (0.012)	0.238* (0.017)
Postmatric	0.074 (0.008)	0.078 (0.009)
Labour force participation:	0.603 (0.012)	0.738* (0.015)
Household composition and resources:		
Household size	6.881 (0.092)	2.487* (0.059)
Lives with mother (grandmother)	0.494 (0.013)	0.129* (0.011)
Lives with father (grandfather)	0.228 (0.011)	0.064* (0.009)
Household includes at least one child with an absent mother	0.462 (0.014)	-
Connected to piped water	0.579 (0.015)	0.655* (0.020)
Connected to electricity	0.747 (0.014)	0.704 (0.018)
Total monthly earnings per adult equivalent	311.131 (18.802)	799.327* (47.023)
Number co-resident in household:		
Economically inactive women	0.714 (0.026)	0.147* (0.013)
Economically inactive men over 18	0.472 (0.022)	0.229* (0.024)
Employed men over 18	0.399 (0.019)	0.478 (0.027)
Pension-aged men	0.102 (0.008)	0.042* (0.008)
Pension-aged women	0.256 (0.012)	0.078* (0.009)
Other children 0 to 18 (resident)	2.758 (0.055)	-
N	2,613	1,448

Notes: The data are weighted and control for clustering and stratification in sample design. Sample sizes are unweighted. Standard errors are in parentheses. The sample includes African non-mothers aged 20 to 49 who are not currently in school. Children are 18 years or younger. Earnings are in 2002 prices and do not include non-earned income. "Total monthly earnings" are normalised by the adult equivalent size of the household which is constructed by signing a value of one to individuals over the age of 11 while children under the age of 12 are assigned a value of 0.5. * indicates that means/ proportions are significantly different across the two groups of non-mothers using a 95 percent confidence interval

Table 4.10 illustrates that there are significant differences in average characteristics across non-mothers who live and do not live with other children. About 56 percent of non-mothers who are co-resident with other children were daughters or granddaughters of the household head, compared to 16 percent of non-mothers not living with any children. Non-mothers who are co-resident with children are significantly younger (29.5 years compared to 33.8 years) and they are less likely to be married. They are also significantly more likely to be living in rural areas and in much larger households. Years of schooling do not differ significantly across the two groups of non-mothers; however non-mothers who are co-resident with other children are significantly more likely to have a matric education. An interesting finding in the table is that non-mothers who live with children reside in households with significantly lower levels of earned monthly income per adult equivalent (R311 compared to R799).

There are also significant differences in labour force participation rates across the two groups of non-mothers. Among non-mothers who are co-resident with other children, sixty percent are labour force participants compared to 74 percent of non-mothers who are not co-resident with other children. This result provides initial evidence that living with other children may impose constraints on the labour force participation of non-mothers who may provide informal childcare within the household.

Conclusion

In the first section of this chapter I described motherhood status among African women aged 20 to 49. In 2002, about 72 percent of African women aged 20 to 49 were mothers, defined as women with biological children 18 years or younger, regardless of co-residency with children. The remaining 28 percent were non-mothers. Findings from the GHS 2002 show marked differences in the marital status, educational attainment and household living arrangements across these two groups of women. Mothers are more likely to be married and living in rural areas and are less likely to be living with their own parents (grandparents to their children); and consistent with international studies, mothers have less education than non-mothers (Budig and England 2001). An age cohort study of completed years of

schooling suggests that educational attainment has increased more among non-mothers than mothers in post-apartheid South Africa. This suggests that the opportunity cost of having children in terms of forgone education has increased in recent years among African women.

The chapter has highlighted further that socio-economic characteristics, especially labour force participation rates, differ markedly not only across mothers and non-mothers, but also according to their co-residency with children. A sizeable sample of all mothers is not co-resident with their children. In 2002, 15 percent of all African mothers were not co-resident with any of their children aged 18 years or younger. These women are probably labour migrants who have left their children in the care of others in the household to work or to search for work (Posel and van der Stoep 2008). The remaining 85 percent of mothers were co-resident with at least one of their children. Compared to not co-resident mothers, co-resident mothers are more likely to be married and living in rural areas. They also face markedly different living arrangements; in particular, they live in larger households which may include children who are not their own.

Labour force participation rates were highest among the group of not co-resident mothers and lowest among co-resident mothers. The analysis indicated that higher labour participation rates among not co-resident mothers, compared to co-resident mothers, appear to be correlated with differences in observed characteristics. The descriptive analysis illustrated that not co-resident mothers have on average fewer biological children who are older; specifically, they are significantly less likely to have children younger than four years of age. Fewer and older children would require less direct care from a mother, facilitating labour migration to places of employment.

Among co-resident mothers, specifically, labour force participation was correlated with the age of the mother's youngest child. Labour force participation rates were lowest among the group with children under four years of age. But there was a more than ten percentage point increase in labour force participation rates when children reached crèche age of four to six years. These results suggest that living with very young biological children particularly constrains women's labour force participation.

Finally, I compared the characteristics of non-mothers distinguished by whether or not they live with other children. Among non-mothers, the majority were co-resident with other children in the household. This group of non-mothers had significantly lower labour force participation rates than non-mothers living without children. This provides initial evidence that the presence of children constrains not only the labour market activity of mothers but of non-mothers living with children who are not their own.

Appendix 4A: Motherhood and co-residency with children

Table 4.A. 1: Motherhood and co-residency with children, women and children 2002

	Among all women aged 20 to 49				Among co-resident mothers	Among all children
	Mother, co-resident with own children	Mother, not co-resident with own children	Not mother, co-resident with other children	Not mother, not co-resident with other children	Co-resident with own and not own children	Not co-resident with mother
All						
Proportion	0.617 (0.004)	0.093 (0.003)	0.169 (0.003)	0.121 (0.004)	0.400 (0.007)	0.249 (0.004)
N	6,161,946	925,633	1,693,678	1,211,538	2,467,443	4,741,953
African						
Proportion	0.613 (0.005)	0.108 (0.003)	0.191 (0.004)	0.089 (0.003)	0.463 (0.007)	0.279 (0.005)
N	4,700,690	828,456	1,461,702	682,047	2,178,511	4,427,014
Coloured						
Proportion	0.707 (0.012)	0.056 (0.005)	0.136 (0.009)	0.102 (0.008)	0.361 (0.019)	0.145 0.009
N	675,348	53,312	130,045	97,039	244,053	236,772
Indian						
Proportion	0.626 (0.023)	0.029 (0.007)	0.134 (0.018)	0.211 (0.024)	0.111 (0.027)	0.048 (0.009)
N	172,309	8,052	36,898	58,088	19,144	17,064
White						
Proportion	0.565 (0.018)	0.032 (0.005)	0.058 (0.007)	0.345 (0.018)	0.041 (0.008)	0.051 (0.009)
N	610,591	34,839	62,545	373,241	24,860	61,103

Source: GHS 2002.

Notes: The data are weighted and adjust for clustering and stratification. Counts are weighted. Standard errors are in parentheses. The data for mothers and non-mothers are for women aged 20-49 and include women who are currently in school. Children are aged 18 years or younger.

The table above highlights racial differences in the living arrangements of women and children specifically. Consistent with the results in Table 3.7, the proportion of women who are mothers not co-resident with any of their own children is highest among African women followed by Coloureds, and lowest among Indians and Whites. Analogously a

larger proportion of African children aged 18 years or younger reported not living with their mothers compared to other race groups. In 2002 about 28 percent (or 4.4 million) of all African children, were not living with their mothers compared to 15 percent of Coloured children and five percent of Indian or White children.

Compared to other race groups, African mothers co-resident with at least one own child are also more likely to be living with other children in the same household. Almost half of African co-resident mothers live also with children who are not their own compared to 36 percent for the Coloured sample, 11 percent for the Indian sample and four percent for the White sample. African non-mothers are also significantly more likely to be living with other children: seventy percent of African non-mothers were living with other children compared to 57 percent of Coloured non-mothers, 39 percent of Indian non-mothers and 14 percent of White non-mothers.[29]

[29] 70 percent ≈[1,461,702 / (1,461,702 + 682,047)]*100; 57 percent ≈ [130,045 / (130,045 + 97,039)]*100; 39 percent ≈ [36,898 /(36,898 + 58,088)]*100; 14 percent ≈ [62,545 /(62,545 + 373,241)]*100

Appendix 4B: Labour force participation rates of African non-mothers and mothers

Table 4.A. 2: Labour force participation rates of African non-mothers and mothers , by age cohort, 2002

	20 – 24	25 – 29	30 - 34	35 – 39	40 – 44	45 – 49
Non-mother	0.611	0.708	0.665	0.676	0.655	0.618
	(0.016)	(0.019)	(0.028)	(0.028)	(0.027)	(0.027)
Mother: co-resident with at least one (own) child	0.454	0.603	0.626	0.670	0.666	0.590
	(0.017)	(0.016)	(0.013)	(0.013)	(0.015)	(0.019)
Mother: not co-resident with (own) children	0.664	0.778	0.797	0.823	0.757	0.764
	(0.032)	(0.022)	(0.024)	(0.028)	(0.033)	(0.041)

Source: GHS 2002

Notes: The data are weighted and control for clustering and stratification in sample design. Unweighted sample sizes of non-mothers, co-resident mothers and not co-resident mothers are 4,061; 9,488 and 1,876 respectively. Standard errors are in parentheses. The sample includes African women aged 20 to 49 who are not currently in school.

Appendix 4C: Age of a mother's youngest children

Table 4.A. 3: The proportion of mothers whose youngest child is of a specific age, African mothers aged 20 to 49, 2002

Age of youngest child	Mothers: co-residing with at least one (own) child	Mothers: not co-resident with (own) children
0	0.144 (0.005)	0.056 (0.006)
1	0.115 (0.005)	0.048 (0.006)
2	0.109 (0.004)	0.079 (0.008)
3	0.101 (0.004)	0.092 (0.008)
4	0.094 (0.004)	0.081 (0.008)
5	0.068 (0.003)	0.088 (0.008)
6	0.064 (0.003)	0.069 (0.007)
7	0.053 (0.003)	0.073 (0.007)
8	0.050 (0.003)	0.064 (0.007)
9	0.041 (0.002)	0.057 (0.007)
10	0.034 (0.002)	0.053 (0.006)
11	0.027 (0.002)	0.035 (0.005)
12	0.024 (0.002)	0.041 (0.006)
13	0.018 (0.001)	0.037 (0.005)
14	0.015 (0.001)	0.031 (0.004)
15	0.014 (0.001)	0.023 (0.004)
16	0.011 (0.001)	0.033 (0.004)
17	0.009 (0.001)	0.021 (0.005)
18	0.008 (0.000)	0.020 (0.004)
N	9,488	1,876

Source: GHS 2002

Notes: The data are weighted and control for clustering and stratification in sample design. Standard errors are in parentheses. Each column of proportions sums to one. The sample includes African women aged 20 to 49 who are not currently in school.

Chapter Five:
Estimating the relationship between motherhood and female labour force participation

Introduction

In this chapter I directly address the main objective of this thesis: to estimate the effects of motherhood on women's labour force participation in South Africa. Recalling the discussion in Chapter One, empirical studies in developed countries have generally found a negative relationship between fertility and labour force participation (cf. Boushey 2008; Chapman *et al* 2001; Angrist and Evans 1998; Bronars and Grogger 1994). In developing countries, however, there is less conclusive evidence of a negative child effect (cf. Aguero and Marks 2008; Posel and van der Stoep 2008; Vere 2008; Cruces and Galiani 2007; Bianchi 2000; Tiefenthaler 1997; Gormick *et al* 1996). I use the General Household Survey 2002 to investigate the relationship between childbearing and labour force participation in South Africa for a sample of African women aged 20 to 49.

A key problem, however, is that identifying the cause and strength of this relationship is complicated by the endogeneity of fertility. Decisions to participate in the labour force and to bear children are likely to be jointly determined. Additionally, there may be omitted factors (or unobserved heterogeneity) that simultaneously influence the decision to have children and the decision to participate in the labour force (Aguero and Marks 2008; Posel and van der Stoep 2008). In this chapter, I do not control directly for this endogeneity using instrumental variable estimation. Rather the objective here is to show how the *definition* of motherhood can exacerbate sources of endogeneity bias in labour force estimations (Posel and van der Stoep 2008). In the next chapter, I evaluate possible instruments for motherhood/childbearing status, given available data in South Africa, and I consider their implications for estimating the relationship between motherhood and female labour force participation.

Empirical studies have typically defined motherhood in terms of co-residency with children. There are three possible reasons for the restriction on co-residency. First, birth history information is often not available in household surveys; therefore, household relationships questions are used to match women to children who live in the same household (Vere 2008; Angrist and Evans 1998; Bronars and Grogger 1994). Second, few mothers are expected to live apart from their children aged 18 years or younger (Vere 2008). Third, there is little reason to expect children to encumber the labour force participation of not co-resident mothers who are not involved in day-to-day childrearing activities. Brewster and Rindfuss (2000:273) note that "it is not childbearing per se but child rearing, the process of caring for and raising a child from birth to adulthood, that leads to the negative relationship between fertility and labour force participation".

In South Africa, however, a sizeable sample of African mothers does not live with any of their children. This was illustrated in the descriptive results of Chapter Four. In 2002, 15 percent of African mothers were not co-resident with any of their own children (see Table 4.1 in Chapter Four). A key reason for absence of mothers is labour migration to places of employment where women leave their children behind in the household of origin to work or to find work. In Chapter Four I showed that these absent or not co-resident mothers are different from mothers who remain co-resident with their children in terms of their observed characteristics. Compared to not co-resident mothers, co-resident mothers are more likely to be married and living in rural areas and they have significantly who are on average older than children of not co-resident mothers. Furthermore they have lower labour force participation rates. These two groups of mothers may also be different in terms of their unobserved characteristics. For example, mothers who are co-resident with children may be a distinct sample of women with less motivation to participate in the labour force. They may also face very different childcare constraints. Mothers who are not co-resident with their children may face less direct childcare responsibilities and therefore have more opportunity to pursue employment opportunities than co-resident mothers. I expect that a co-residency requirement on motherhood would over-estimate the negative relationship between motherhood and labour force participation (Posel and van der Stoep 2008). I test

this by removing co-residency restrictions on motherhood definitions in labour force estimations.

If a considerable sample of African mothers live apart from their children, then other women in the child's household are likely to act as maternal surrogates. This is consistent with complex family structures and extended living arrangements in African households where the fosterage of children is a common occurrence, particularly in rural areas (Posel and van der Stoep 2008; Amoateng *et al* 2007; Anderson and Phillips 2006; McDaniel and Zulu 1996; Van der Waal 1996). A significant proportion of African women in the GHS 2002 sample are living with children who are not biologically their own (see Table 4.2 in Chapter Four). About 46 percent of co-resident mothers, and almost seventy percent of non-mothers, live in households with other children. These estimates are particularly pronounced in rural areas at 52 percent and 73 percent for co-resident mothers and non-mothers respectively. Furthermore, among African non-mothers living with children aged 20 to 49, 46 percent were living with a child whose biological mother was absent from the household in 2002. I use these sample characteristics as an opportunity to investigate and compare the effects on women's labour force participation of living with biological children versus other children. In particular, I consider whether caring for children constrains the labour force participation of non-mothers.

In the next section I outline model specifications for estimating the relationship between motherhood and labour force participation. In section 5.2 I present regression results where mothers are identified as co-resident with children but in section 5.3 this co-residency requirement is removed.

5.1 Model specification

Where motherhood has been treated as exogenous within labour force estimations, empirical studies have generally used ordinary least squares (OLS) models (Aguero and Marks 2008) and occasionally probit models (Boushey 2008) to estimate the relationship between motherhood and labour force participation. Estimation results throughout this

chapter are robust to whichever of the two models are used. However, I have chosen to present results from the OLS[30] model because the estimated coefficients are more easily interpreted than in the probit model.

The OLS model is represented as follows:

$$Yi = \beta Mi. + \gamma Ii + \delta H_i + e_i$$

Y_i is a dependent variable taking on a value of one if a woman is labour force participant and zero if she is economically inactive. Consistent with Chapter Four I use a strict definition of labour force participation, defining a woman as a labour force participant if she was either employed or unemployed but actively searching for work. M_i represents motherhood status, I_i is a vector of individual characteristics, H_i is a vector of household characteristics and e_i is a random error term. β is the coefficient of interest and represents the effect of childbearing/motherhood on women's labour force participation.

Determining the "true" motherhood effect is compromised by endogeneity bias where motherhood status M_i is expected to be correlated with the error term, e_i. There are two reasons for this. First, a two-way causal relationship may exist between motherhood and labour force participation. Second, omitted variables such as ambition or motivation, captured in the error term, e_i, may be correlated with both the decision to enter motherhood and the decision to participate in the labour force. If ambition or motivation is lower among mothers than non-mothers, and positively correlated with labour force participation, then the size of the negative motherhood effect may be upwardly biased; β will capture both the effect of motherhood status and indirect effects of unobserved characteristics on labour force participation. Although I do not explicitly control for endogeneity bias in labour force estimations in this chapter, I show that the definition of motherhood can exacerbate sources of endogeneity bias when estimating the motherhood effect, β. I show this by estimating and comparing two sets of regressions distinguished by how mothers are identified. The

[30] Results from probit estimations for all the regressions considered in this chapter are reported in Tables 5A.1-4 in Appendix 5A.

first set of regressions imposes a co-residency requirement when identifying mothers. The second set of regressions removes the co-residency requirement.

Considering the first set of regressions, I run three model specifications. The first specification includes a dummy variable which takes on a value of one if a woman has at least one biological child aged zero to 18 years and is *co-resident* with at least one of her children. The reference category here is non-mothers and women who have biological children aged 18 years of younger but who do not reside with any of these children. The second specification includes four dummy variables for the age of the mother's youngest child. The descriptive results in Chapter Four suggest that labour force participation among co-resident mothers varies by the age of her children. Very low labour force participation rates among mothers with children younger than four years suggest that labour force participation is particularly constrained among mothers with very young children. I now test econometrically whether this result holds, controlling for other observed characteristics. Specification three recognises that living with children who are not biologically one's own may also have negative effects on female labour force participation. I include two binary variables for "own" child and "other" child as well as their interaction term (O*N). When "own" child equals zero and "other" child is one, then "other" child identifies non-mothers who live in a household with children. If both "own" child and "other" child are equal to one, the interaction term will identify co-resident mothers who are also living with children who are not their own.

In the second set of regressions I remove the co-residency requirement in identifying mothers. By including two binary variables to indicate whether a woman is a co-resident mother or a not co-resident mother, I compare how the effects of motherhood on labour force participation vary by co-residency with children. I then combine these two groups of mothers into one category, identifying mothers as women who have children aged 18 years or younger, regardless of co-residency status.

I then rerun the first and second sets of regressions for a rural sample of African women only and compare these results to the national sample. Where migration occurs particularly

from rural to urban areas, mothers who remain co-resident with their children in rural areas are likely to be a particularly non-random sample of all mothers. The negative relationship between motherhood and labour force participation is therefore expected to be particularly biased by a co-residency requirement in rural areas. I test this by restricting the sample to rural women only and comparing motherhood estimates across the national and rural samples.

In every regression specification I control for a range of other individual and household covariates including the woman's age and age squared, her educational attainment and marital status, rural/urban status, nine dummies for province of residence and four household composition variables. In developing country studies it is also important to recognise that women's ability to participate in market activities is linked to the availability of infrastructure such as piped water and electricity (Desai and Jain 1994). Collecting firewood and water, for example, may impose time constraints on a woman's ability to participate in the labour force. I therefore also include two dummy variables for whether a woman lives in a household with access to piped water and electricity. Means and standard errors of these indicators are presented in Table 4.5 and Table 4.6 in Chapter Four.

5.2 Labour force participation among co-resident mothers

I present results of the first three specifications in Table 5.1. Here motherhood is identified in terms of co-residency with children. In the first specification (I) the negative and significant coefficient on the motherhood dummy variable shows that relative to non-mothers and not co-resident mothers, co-resident mothers are significantly less likely to be labour force participants. Where developed and developing country studies have used OLS or probit estimations (i.e. have not explicitly controlled for the endogeneity of motherhood), the motherhood effect has been estimated between negative nine and negative 14 percent since the late 1990's (Aguero and Marks 2008; Boushey 2008). The size of the negative motherhood effect in South Africa is comparatively smaller at about six percent.

Table 5. 1: Co-resident mothers and labour force participation, African women 2002

	I	II	III
Mother	-0.063***		
	(0.046)		
Children 0-3		-0.116***	
		(0.011)	
Children 4-6		-0.026*	
		(0.015)	
Children 7-12		-0.005	
		(0.013)	
Children 13-18		-0.004	
		(0.020)	
Own children 0-18 (O)			-0.118***
			(0.013)
Not own children 0-18 (N)			-0.067***
			(0.015)
O*N			0.100***
			(0.018)
Age	0.046***	0.044***	0.046***
	(0.005)	(0.005)	(0.005)
$Age^2/100$	-0.057***	-0.057***	-0.056***
	(0.007)	(0.007)	(0.007)
Primary	0.122***	0.119***	0.120***
	(0.019)	(0.019)	(0.019)
Incomplete secondary	0.162***	0.158***	0.160***
	(0.020)	(0.019)	(0.019)
Matric	0.246***	0.242***	0.246***
	(0.021)	(0.020)	(0.020)
Diploma/degree	0.347***	0.343***	0.348***
	(0.021)	(0.021)	(0.021)
Married	-0.128***	-0.119***	-0.124***
	(0.011)	(0.011)	(0.012)
Divorced/widowed	0.005	0.003	0.007
	(0.017)	(0.017)	(0.017
Rural	-0.077***	-0.074***	-0.076***
	(0.014)	(0.014)	(0.014)
Piped water	0.035***	0.034**	0.035***
	(0.013)	(0.013)	(0.013)
Electricity	0.040***	0.038***	0.040***
	(0.002)	(0.013)	(0.013)
F	86.13	81.55	82.40
N	15,425	15,425	15,425

Source: GHS 2002.

Notes: The sample is all women aged 20-49. Mothers exclude all biological mothers who are not co-resident with at least one of their own children. The data are weighted. The regressions control for stratification and clustering in sample design. Standard errors are in parentheses. The omitted categories are no schooling and never married. The regressions also include nine dummy variables for province of residence and four controls for household composition (the number of not economically active men and women; and the number of male and female pensioners). *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Consistent with international studies (Aguero and Marks 2008; Boushey 2008; Chapman *et al* 2001; Gormick *et al* 1996), the results in the second specification (II) show that the negative relationship between motherhood and labour force participation derives particularly from living with very young children. Living with own children under four years reduces women's labour force participation by 12 percent. However, when children reach crèche age of four to six years the negative child effect is only three percent and weakly significant. At school-going age, children do not impose a significant negative effect on a mother's labour force participation (Posel and van der Stoep 2008).

The third specification (III) isolates the effects of living with own and other children on women's labour force participation. The three coefficients are interpreted relative to the reference category which is mothers and non-mothers who are not living with any children. The coefficient on own child is negative and significant (-0.118). Mothers who co-reside *only* with *own* children are significantly less likely to be labour force participants. Living with other children who are not biologically your own also has a negative effect on the labour force participation of non-mothers. Although the size of this effect is smaller than the size of the effect of living with own children (-0.067 compared to -0.118), the results show that the costs of childrearing in South Africa are not borne by mothers alone. The coefficient on the interaction term is positive and significant. Because the magnitude of the coefficient on the interaction term is larger than the magnitude of the coefficient on other child (0.1 compare to -0.067), the overall effect of living with other children, in addition to own children, does not strengthen the negative relationship between motherhood and labour force participation (*ibid*).

Across the first sets of regressions results, coefficient estimates on household and other individual explanatory variables consistently have the same signs and similar magnitudes. Coefficients on age and educational dummies are positive and significant. Therefore labour force participation increases non-linearly with both a woman's age and educational status. Women who are married and live in rural areas are significantly less likely to be labour force participants. Living in households with access to piped water and electricity positively affects women's labour force participation. This indicates that the provision of

services such as piped water and electricity frees up time in household production (such as collecting water or wood for fuel) which is reallocated to market work.[31]

5.3 Labour force participation among all mothers

In the previous set of regressions, a co-residency requirement was imposed upon mothers. In this second set of regressions, the definition of motherhood is expanded to include a large sample of mothers who are not co-resident with their children.

In the fourth specification I distinguish between co-resident and not co-resident mothers by including two dummy variables. The reference category now represents only non-mothers. Results are presented in Table 5.2. The coefficient on both co-resident mother and not co-resident mother are both significant but have opposite signs. Compared to non-mothers, not co-resident mothers are significantly more likely to be labour force participants. An important observation here is that the *co-resident* motherhood effect is now about negative four percent compared to negative six percent in specification I. The reason for this difference is that in specification I, which imposed a co-residency requirement on motherhood, included in the reference category of 'non-mothers' were mothers who are not co-resident with their children. But not co-resident mothers are (even) more likely than non-mothers to participate in the labour force. The negative effect of being a co-resident mother on women's labour force participation is therefore overestimated if not co-resident mothers are included in the reference category.

In the fifth specification I combine both co-resident and not co-resident mothers into a single binary variable which takes on a value of one if a woman has at least one biological child aged 18 years or younger, regardless of co-residency with children. The size of the

[31] The piped water and electricity dummies could also be argued to be proxies for income levels in the household. In this case the signs are expected to be negative because in the face of higher income levels, there is less economic need for women to participate in the labour force. But the causality could run in the opposite direction – women who work are more able to afford these services which may explain the positive relationship.

negative, 'aggregate' motherhood effect is now only two percent and only weakly significant (see specification V). The size of the 'aggregate' motherhood effect is considerably smaller than the size of the co-resident motherhood effect (see specification I in Table 5.1). This confirms that a co-residency requirement on motherhood upwardly biases the negative relationship between motherhood and labour force participation in South Africa.

Table 5. 2: All mothers and labour force participation, African women 2002

	IV	V
Co-resident mother	-0.038***	
	(0.011)	
Not co-resident mother	0.074***	
	(0.015)	
All mothers		-0.019*
		(0.010)
Age	0.043***	0.043***
	(0.005)	(0.005)
$Age^2/100$	-0.052***	-0.052***
	(0.007)	(0.007)
Primary	0.120***	0.119***
	(0.019)	(0.019)
Incomplete secondary	0.159***	0.158***
	(0.019)	(0.020)
Matric	0.246***	0.245***
	(0.020)	(0.021)
Diploma/degree	0.347***	0.344***
	(0.021)	(0.021)
Married	-0.130***	-0.139***
	(0.011)	(0.011)
Divorced/widowed	0.004	-0.002
	(0.017)	(0.017)
Rural	-0.077***	-0.079***
	(0.014)	(0.014)
Piped water	0.035***	0.036***
	(0.013)	(0.013)
Electricity	0.041***	0.037***
	(0.013)	(0.013)
F	86.08	81.29
N	15,425	15,425

Source: GHS 2002.

Notes: The sample is all women aged 20-49. The data are weighted. The regressions control for stratification and clustering in sample design. Standard errors are in parentheses. The omitted categories are no schooling and never married. The regressions also include nine dummy variables for province of residence and four controls for household composition including the number of not economically active women aged 19 to 59, the number of not economically active men aged 19 to 64, the number of male pensioners (over 64 years) and the number of female pensioners (over 59 years). *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table 5. 3: Labour force participation among African mothers, 2002

	No control for age of children	Include control for age of children
Not co-resident mother	0.113***	0.088***
	(0.013)	(0.013)
Age of youngest child		0.010***
		(0.001)
Age	0.045***	0.045***
	(0.006)	(0.006)
$Age^2/100$	-0.054***	-0.057***
	(0.008)	(0.008)
Primary	0.102***	0.101***
	(0.022)	(0.022)
Incomplete secondary	0.142***	0.137***
	(0.023)	(0.023)
Matric	0.225***	0.223***
	(0.024)	(0.024)
Diploma/degree	0.332***	0.331***
	(0.025)	(0.025)
Married	-0.139***	-0.131***
	(0.013)	(0.013)
Divorced/widowed	-0.018	-0.023
	(0.019)	(0.019)
Rural	-0.085***	-0.081***
	(0.016)	(0.016)
Piped water	0.046***	0.044***
	(0.014)	(0.014)
Electricity	0.033**	0.030***
	(0.014)	(0.014)
F	71.96	74.30
N	11,364	11,364

Source: GHS 2002.

Notes: The sample is all *mothers (regardless of residency with children)* aged 20 to 49. A mother is a woman with biological children aged 18 years or younger. The reference category is co-resident mothers. The data are weighted. The regressions control for stratification and clustering in sample design. Standard errors are in parentheses. The omitted categories are no schooling and never married. The regressions include the same controls as in Table 5.1 and 5.2. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

The analysis above has highlighted that not co-resident mothers are about nine percent more likely to be labour force participants compared to co-resident mothers, despite controlling for differences in the age of their children. A possible reason for this result is that not co-resident mothers do not face direct childcare constraints and therefore have more opportunity to work or to search for work (Posel and van der Stoep 2008). However a reverse causal relationship may exist between labour force participation and co-residency

with children. Women who want or need to work may be less likely to live with their children (*ibid*; Van der Waal 1996).[32] I am however unable to control for the endogeneity of mother's living arrangements with their children with respect to their labour force participation.[33]

An important part of labour force participation in South Africa is migration to places of employment. In this case, women leave their children in the care of others in the household of origin to work or to search for work (Posel and van der Stoep 2008). Labour migration occurs particularly from rural areas. For this reason the remaining sample of co-resident mothers in rural areas is likely to be a particularly non-random sample of all mothers. I expect a larger negative motherhood effect in rural areas where motherhood status is particularly endogenous with respect to labour force participation (*ibid*). I test this by rerunning the five earlier regressions (reflected in Tables 5.1 and 5.2) on a sample of only *rural* African women aged 20 to 49. The rural sample results are reflected in Table 5.4.

There is a significant negative motherhood effect of nine percent when a co-residency restriction is placed on motherhood; and compared to the national sample the rural sample effect is three percentage points larger. Rural co-resident mothers are less likely to be labour force participants even when their children mature in age. The child age dummies are now negative, significant and larger in magnitude up to age 13 (Specification II in Table 5.4) This contrasts with the national results where only very young children (under four years of age) have a significant and sizable negative effect on women's labour force participation (Specification II in Table 5.1). Living with *own* children *only* continues to have a large and significant negative effect on women's labour force participation in rural

[32] For example Van der Waal (1996) finds that a primary reasoning that African mothers leave their children behind in rural areas is due to an inability to financially support their children. Mothers are forced to leave children at home to look for work elsewhere.

[33] There are no available variables which would be suitable instruments for living arrangements i.e. variables that are correlated with a woman's living arrangements but uncorrelated with her decision to participate in the labour force. Although the age of a woman's children may be correlated with her co-residency status, the results in Table 5.3 indicate that the age of a woman's child is not independent of her labour force participation.

areas. Non-mothers are also significantly less likely to be labour force participants if they live with other children; but the size of this effect is about four percentage points larger than in the national sample (-0.105 in Table 5.4 compared to -0.067 in Table 5.1).

The results in specification IV for the rural sample also show that relative to non-mothers and co-resident mothers, not co-resident mothers in rural areas are significantly more likely to be labour force participants. Interestingly, the coefficient on not co-resident mother is larger in the rural sample estimation than in the national estimation (0.111 in Table 5.4 compared to 0.074 in Table 5.2). This suggests that labour migration is not only occurring between rural and urban areas but also within rural areas (*ibid*).

Table 5. 4: Motherhood and labour force participation, *rural* African women 2002

	I	II	III	IV	V
Mother (co-resident)	-0.094***				
	(0.014)				
Children 0-3		-0.138***			
		(0.017)			
Children 4-6		-0.054**			
		(0.021)			
Children 7-12		-0.041**			
		(0.020)			
Children 13-18		-0.036			
		(0.032)			
Own children 0-18 (O)			-0.189***		
			(0.022)		
Not own children 0-18 (N)			-0.105***		
			(0.025)		
O*N			0.148***		
			(0.028)		
Mother (co-resident)				-0.057***	
				(0.017)	
Mother (not co-resident)				0.111***	
				(0.023)	
Mother (no residency requirement)					-0.033**
					(0.016)

Source: GHS 2002.

Notes: The sample is all women aged 20-49. The data are weighted. Standard errors are in parentheses. The same set of other independent variables in Table 5.1 and 5.2 are included in these regressions. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

The coefficient on the combined motherhood variable (which includes both co-resident and not co-resident mothers into one category) is negative and significant in specification V, Table 5.4. However the size of the 'aggregate' motherhood effect for the rural sample drops by six percentage points, compared to a fall of four percentage points for the national sample, when a co-residency requirement is removed (compare specification V to specification I). This suggests that imposing a co-residency requirement on motherhood exaggerates sources of endogeneity bias in the relationship between labour force participation and motherhood, particularly in rural areas (*ibid*).

Conclusion

International studies that estimate the effect of motherhood on women's labour force participation have typically defined mothers as co-resident with their children in the household. If mothers are identified as women with at least one co-resident biological child, I find evidence of a significant negative relationship between motherhood and labour force participation for a sample of African women aged 20 to 49; and this negative relationship is driven particularly by the presence of very young children in the household.

The size of the negative motherhood effect in South Africa is about six percent if a co-residency restriction is placed on motherhood. This is comparatively smaller than observed motherhood effects in other countries when studies do not control for the endogeneity of motherhood in labour force estimations (Aguero and Marks 2008; Boushey 2008). However the size of the motherhood effect in South Africa is biased upward for two reasons. First, the coefficient on the motherhood variable may be capturing indirect effects of omitted variables such as ability or motivation which positively affect labour force participation. If mothers are less motivated or ambitious than non-mothers in their careers, this unobserved difference may exaggerate the true effect of motherhood on women's labour force participation. Second, the effect may be upwardly biased because co-resident mothers are a select sample of all mothers who are less likely to be labour participants. Although I was not able to address the omitted variable bias, this chapter shows that the

second source of bias is particularly present in estimating the relationship between motherhood and women's labour force participation in South Africa.

A sizeable sample of all African mothers in South Africa does not live with their children: 15 percent were not co-resident with any of their children aged 18 years or younger in 2002. Not co-resident mothers are probably labour migrants who have left children in the household of origin. Compared to non-mothers and co-resident mothers, not co-resident mothers are significantly more likely to participate in the labour force. Expanding the definition of motherhood to include mothers who are not co-resident with their children reduces the size of the negative motherhood effect from six percent to two percent (a fall of four percentage points). In rural areas specifically, the definition of motherhood makes particularly large differences to the estimated motherhood effect.

Given that a large sample of African mothers lives apart from their children, childcare is probably provided by other women in the household of origin. In 2002, about 46 percent of co-resident mothers and almost seventy percent of non-mothers lived in a household with other children. This chapter explored whether surrogate parenthood or fostering of children impacted on the labour force participation of co-resident mothers and non-mothers who live with children who are not their own. I find that non-mothers who live with other children are less likely to be labour force participants; but the size of this effect is smaller than the co-resident motherhood effect. Among mothers with own children in the household, living with other children as well does not strengthen the negative relationship between motherhood and labour force participation.

In this chapter I did not directly control for possible endogeneity bias in the relationship between motherhood and labour force participation. Therefore OLS estimates of the size of the motherhood effect are probably overestimated in the presence of omitted variables. Despite this possible bias, when a co-residency requirement is removed from the definition of motherhood, the size of negative 'aggregate' motherhood effect among African women in South Africa is only two percent and weakly significant.

Appendix 5A: Probit estimations

Table 5.A. 1: Probit estimation corresponding to Table 5.1 - co-resident mothers and labour force participation, African women 2002

	I	II	III
Mother	-0.188***		
	(0.029)		
Children 0-3		-0.338***	
		(0.032)	
Children 4-6		-0.080*	
		(0.045)	
Children 7-12		-0.013	
		(0.041)	
Children 13-18		-0.008	
		(0.062)	
Own children 0-18 (O)			-0.375***
			(0.043)
Not own children 0-18 (N)			-0.234***
			(0.051)
O*N			0.330***
			(0.058)
Age	0.137***	0.133***	0.136***
	(0.014)	(0.014)	(0.014)
$Age^2/100$	-0.169***	-0.171***	-0.168***
	(0.021)	(0.021)	(0.021)
Primary	0.335***	0.327***	0.333***
	(0.052)	(0.052)	(0.052)
Incomplete secondary	0.452***	0.440***	0.447***
	(0.054)	(0.054)	(0.054)
Matric	0.708***	0.699***	0.711***
	(0.059)	(0.059)	(0.059)
Diploma/degree	1.144***	1.134***	1.153***
	(0.077)	(0.077)	(0.077)
Married	-0.389***	-0.366***	-0.382***
	(0.036)	(0.036)	(0.036)
Divorced/widowed	0.024	0.015	0.028
	(0.057)	(0.057)	(0.042)
Rural	-0.230***	-0.221***	-0.227***
	(0.042)	(0.042)	(0.042)
Piped water	0.103***	0.100***	0.105***
	(0.039)	(0.039)	(0.038)
Electricity	0.137***	0.104***	0.113***
	(0.037)	(0.037)	(0.037)
F	53.20	49.87	50.58
N	15,425	15,425	15,425

Source: GHS 2002.

Notes: The sample includes all African women aged 20-49. Mothers exclude all biological mothers who are not co-resident with at least one of their own children. The data are weighted. The regressions control for stratification and clustering in sample design. Standard errors are in parentheses. The omitted categories are no schooling and never married. The same set of other independent variables in Table 5.1 and 5.2 are included in these regressions. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table 5.A. 2: Probit estimations corresponding to Table 5.2 - all mothers and labour force participation, African women 2002

	IV	V
Co-resident mother	-0.106**	
	(0.032)	
Not co-resident mother	0.258***	
	(0.050)	
All mothers		-0.051
		(0.031)
Age	0.127***	0.126***
	(0.014)	(0.014)
$Age^2/100$	-0.154***	-0.154***
	(0.021)	(0.021)
Primary	0.329***	0.326***
	(0.052)	(0.052)
Incomplete secondary	0.442***	0.438***
	(0.054)	(0.054)
Matric	0.711***	0.703***
	(0.059)	(0.059)
Diploma/degree	1.146***	1.136***
	(0.076)	(0.076)
Married	-0.397***	-0.419***
	(0.035)	(0.035)
Divorced/widowed	0.019	0.002
	(0.057)	(0.057)
Rural	-0.230***	-0.236***
	(0.042)	(0.042)
Piped water	0.105***	0.108***
	(0.039)	(0.039)
Electricity	0.113***	0.102***
	(0.037)	(0.037)
F	52.79	50.93
N	15,425	15,425

Source: GHS 2002.

Notes: The sample includes all African women aged 20-49. The data are weighted. The regressions control for stratification and clustering in sample design. Standard errors are in parentheses. The omitted categories are no schooling and never married. The regressions also include nine dummy variables for province of residence and four controls for household composition including the number of not economically active women aged 19 to 59, the number of not economically active men aged 19 to 64, the number of male pensioners (over 64 years) and the number of female pensioners (over 59 years). *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table 5.A. 3: Probit estimations corresponding to Table 5.3 - labour force participation among African mothers, 2002

	No control for age of children	Include control for age of children
Not co-resident mother	0.370***	0.293***
	(0.045)	(0.046)
Age of youngest child		0.031***
		(0.004)
Age	0.133***	0.133***
	(0.017)	(0.017)
$Age^2/100$	-0.157***	-0.170***
	(0.025)	(0.025)
Primary	0.281***	0.279***
	(0.060)	(0.060)
Incomplete secondary	0.394***	0.384***
	(0.063)	(0.064)
Matric	0.654***	0.656***
	(0.070)	(0.070)
Diploma/degree	1.101***	1.099***
	(0.090)	(0.091)
Married	-0.428***	-0.404***
	(0.040)	(0.040)
Divorced/widowed	-0.047	-0.064
	(0.063)	(0.063)
Rural	-0.256***	-0.246***
	(0.048)	(0.048)
Piped water	0.140***	0.132***
	(0.041)	(0.041)
Electricity	0.089**	0.082**
	(0.041)	(0.040)
F	43.38	44.14
N	11,364	11,364

Source: GHS 2002.

Notes: The sample includes all African *mothers (regardless of residency)* aged 20-49. A mother is a woman with biological children aged 18 years or younger. The reference category is co-resident mothers. The data are weighted. The regressions control for stratification and clustering in sample design. Standard errors are in parentheses. The omitted categories are no schooling and never married. The regressions include the same controls as in Table 5.1 and 5.2. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Table 5.A. 4: Probit estimations corresponding to Table 5.4 - motherhood and labour force participation, *rural* African women 2002

	Ia	II	III	IV	V
Mother (co-resident)	-0.257***				
	(0.040)				
Children 0-3		-0.379***			
		(0.047)			
Children 4-6		-0.148**			
		(0.059)			
Children 7-12		-0.115**			
		(0.055)			
Children 13-18		-0.098			
		(0.089)			
Own children 0-18 (O)			-0.541***		
			(0.066)		
Not own children 0-18 (N)			-0.317***		
			(0.075)		
O*N			0.434***		
			(0.082)		
Mother (co-resident)				-0.151***	
				(0.046)	
Mother (not co-resident)				0.336***	
				(0.070)	
Mother (no residency requirement)					-0.086*
					(0.045)

Source: GHS 2002.

Notes: The sample includes all African women aged 20-49. The data are weighted. Standard errors are in parentheses. The same set of other independent variables in Table 5.1 and 5.2 are included in these regressions. *** Significant at the 1% level; ** Significant at the 5% level; * Significant at the 10% level.

Chapter Six:

Instrumental variable estimation using sibling sex composition

Introduction

The key problem for investigating a causal relationship between childbearing and labour force participation in South Africa is the endogeneity of childbearing. Childbearing and labour force participation decisions may be simultaneously determined; and unobserved heterogeneity in characteristics may be correlated with both of these decisions. Therefore childbearing variables are likely to be correlated with the error term in an equation of labour force participation.

Instrumental variable (IV) estimation has typically been used in the literature in dealing with endogeneity, where instruments are variables which are uncorrelated with the outcome variable (such as labour supply or labour force participation) but highly correlated with the endogenous explanatory variable (for example motherhood or childbearing). Bound *et al* (1995:443) explain that "the exogenous instruments allow the researcher to partition the variance of the endogenous explanatory variable into exogenous and endogenous components. The exogenous component is then used in the estimation". IV estimation therefore allows the researcher to estimate an exogenous effect of childbearing on labour force participation.

In order for IV estimation to yield consistent and unbiased estimates, the exogenous instruments chosen must satisfy two properties. First, the instrument must be uncorrelated with omitted or unobserved variables in the error term. The instrument must therefore have no direct association with the outcome variable of the study, in this case labour force participation. Second, the instrument must be correlated with the endogenous explanatory variable i.e. childbearing (Wooldridge 2003:486; Bound *et al* 1995:443). An instrument that has been argued to satisfy these two properties is the sibling sex composition of a woman's first two children. Same sex sibling composition was first suggested by Angrist and Evans (1998) to investigate the causal effects of further childbearing on mothers'

labour force participation in the United States. It was later used by Iacovou (2001) for a similar study in the United Kingdom. Most recently Cruces and Galiani (2007) have applied this methodology in the developing country context of Argentina and Mexico. Each study argues that whether or not a woman's first two children are of the same sex is expected to be correlated with further childbearing because parents generally have a preference for a mixed sibling sex composition or a 'taste for balance' (Ben-Porath and Welch 1976). But the sex of any given child is argued to be a random occurrence (Clark 2000:95) uncorrelated with background characteristics or unobservable characteristics that influence women's labour force participation. As a result sibling sex composition is expected to be uncorrelated with omitted variables in the error term.

In this Chapter I consider an application of Angrist and Evan's (1998) same sex sibling composition to instrument for further childbearing among African women aged 20 to 49 with at least two children in South Africa. Despite evidence of a strong correlation between sibling sex composition and further childbearing in the United States, the United Kingdom, Argentina and Mexico, there may not be a strong preference for a mixed sibling sex composition among Africans in South Africa. To date there is no empirical evidence to suggest a preference for a mixed sibling sex composition among African women in South Africa given the sex of their first two children (Gangadharan and Maitra 2003; Arnold 1992). A possible reason is that family sizes are larger among Africans in South Africa than among women in the aforementioned countries (Central Intelligence Agency 2008; Ben-Porath and Welch 1976). Among populations with large average family size, parents who have 'taste for balance' may not be concerned about whether their first two children are of the same sex; they eventually expect to obtain a mixed sibling sex composition given desires for larger families (*ibid*:292). Consequently strong correlations between the sex of the first *two* children and further childbearing may not be observed so that same sex sibling composition may not adequately satisfy the second property of an IV in some countries. The implication is that very large standard errors are obtained on IV estimates and coefficients lack precision and any meaningful interpretation (Wooldridge 2003).

In this chapter, therefore, I carefully analyse whether same sex sibling composition satisfies the properties of an instrumental variable for further childbearing among Africans in South Africa. Section 6.1 considers the data and samples used in this analysis and section 6.2 outlines the methodology. I consider the quality of sibling sex composition variables as instruments for further childbearing in South Africa in Section 6.3. Finally, I present second-stage results from the two-stage least squares estimation in section 6.4 and then conclude the chapter.

6.1 Data

Suggested instruments for childbearing in the international literature have included the occurrence of twin births (Bronars and Grogger 1994; Rosenzwieg and Wolpin 1980), women's infertility status (Aguero and Marks 2007) and religious affiliation (Ryder and Westoff 1972 cited in Iacovou 2001). But due to data constraints, these instruments cannot be used for South African studies. Applying twin births as an instrument for childbearing requires very large samples unavailable in South African household surveys because only very few women report multiple births. And information on a woman's infertility status and religious affiliation is typically not collected in nationally representative household surveys. The Chinese Lunar calendar has also been exploited to instrument for childbearing (see Vere 2008), however, its application is limited to countries with dense Chinese populations.

A less data-demanding instrument that has been used is the sibling sex composition of a woman's first two children (Cruces and Galiani 2007; Iacovou 2001 and Angrist and Evans 1998). Information on the sibling sex composition of children born to a woman *is* determinable using the General Household Survey (GHS) 2002. Included among detailed retrospective fertility data on children ever born to a woman in the GHS 2002 is the birth order and sex of each child born alive. It is possible therefore to determine whether the first two children born to a woman are of the same or different sex.

Unfortunately same sex sibling composition can only be used to instrument for whether a woman has more than two children rather than any children at all. The sample of analysis

in this chapter is therefore limited to all African mothers[34] aged 20 to 49 who have had *at least two live births* and are not currently in school. I also continue to exclude all women whose every child has died subsequent to birth. I identify 7,477 African mothers (regardless of co-residency with children) who had experienced at least two live births. About seventy percent of these mothers are married and 87 percent are co-resident with at least one of their children aged 18 years or younger. About 65 percent of the sample of mothers are labour force participants. Consistent with previous chapters I continue to disaggregate the sample according to co-residency with children. I have also further disaggregated the sample by marital status. Compared to single mothers there may be stronger desires among couples for a mixed sibling sex composition. Alternatively, higher fertility rates among married mothers compared to unmarried mothers may lower concerns about the sibling sex composition of their first two children because they eventually expect to obtain a balanced sex composition of offspring. This is further discussed in section 6.3.2. I identify a woman as married if her response to a question on marital status was "married or living together as husband and wife".[35] Unmarried women include women who have never married or are divorced, separated or widowed.

[34] Consistent with previous chapters mothers are defined here as women with at least one biological child aged 18 years or younger. In this case specifically, at least one of the two or more children born to a mother is 18 years or younger.

[35] It is not possible to distinguish women who have had a civil/customary marriage from women who are cohabiting with a partner, but not married, in the GHS 2002. There may be differences in the level of male support and income sharing among couples depending on whether they are married or cohabiting. Furthermore in the event of separation, women in cohabiting partnerships are also less protected by the law and may be more likely to keep their jobs than women in civil marriages as a form or economic security. I acknowledge that these factors may possibly result in differences in labour force participation behaviour by whether women are married or cohabiting with a partner; but I will not be able to observe these differences in this analysis. There may also be heterogeneity in preferences for sibling sex composition among couples who are married compared to cohabiting couples which also cannot be observed in this analysis.

6.2 Methodology

For the sample described above I estimate the following specification:

$$Y_i = \alpha + \beta T_i + \delta X_i + e_i \qquad (1)$$

T_i is the key variable which takes on a value of one if a woman has more than two live births and zero if she only has had two live births. Y_i represents the labour force participation of a mother and is equal to one if a woman is a labour force participant and zero otherwise. X_i is a vector of other explanatory variables including a woman's age, age squared, age at first birth, her educational status and marital status. Household compositional variables, dummy variables for province of residence and whether the woman lives in a rural or urban area are also included among the explanatory variables.

In this estimation, β is supposed to measure the causal relationship between having a third birth and women's labour force participation. Discussed in Chapter One and in the previous chapter, the estimation of β is complicated by the endogeneity of childbearing. First, decisions about childbearing and labour force participation may be determined interdependently. Second, T_i may be correlated with unobserved variables such as ambition or ability in the error term, e_i (Aguero and Marks 2008; Bound *et al* 1995). Ordinary least squares (OLS) estimation of equation (1) gives biased and inconsistent estimates of β, the effect of a third birth on mother's labour force participation. In particular the magnitude of β is likely to be biased upwards due to unobserved ambition or ability which positively influences the outcome variable but is negatively correlated with having a third birth.

In order to address the endogeneity concern in equation (1) I instrument for *more than two births*, T_i. Let S_i be an indicator equal to one if a woman's first two children are of the same sex or zero if the first two children are of mixed sex. Other studies suggest that the variable *more than two children*, T_i, is expected to be positively correlated with same sex sibling composition (Angrist and Evans 1998; Iacovou 2001; Cruces and Galiani 2007). Research indicates that parents generally have a preference for balanced families with both girls and boys (Arnold 1992; Ben-Porath and Welch 1976). In other words women whose first two

children are of the same sex typically are more likely to have third child than women with a mixed sibling sex composition. But there may be different effects on the likelihood of having a third birth depending on whether the first two siblings of the same sex are girls or boys. For example, the effect may be stronger if the first two children born are girls because there may be a stronger preference for sons than daughters. Therefore, I also decompose same sex siblings into two instruments *two girls*, G_i and *two Boys*, B_i which take on a value of one if a woman's first two children born alive were both girls, or both boys, respectively.

Two-stage least squares estimation is then used to obtain the IV estimator of the effect of an exogenous (third) birth on a woman's labour force participation. In the first-stage estimation the endogenous variable T_i is expressed as a linear function of all other exogenous variables X_i in the structural model and of the identifying instrument/s to obtain the reduced form equation. The first-stage equation relating more than two children to sibling sex composition is

$$T_i = \gamma(S_i) + \lambda X_i + \eta_i \qquad (2a)$$

where γ is the first-stage effect of sibling sex composition on the likelihood of having a third birth. Alternatively the first-stage relationship between T_i and sibling sex composition is further decomposed to

$$T_i = \gamma_o(G_i) + \gamma_1(B_i) + \lambda X_i + \eta_i \qquad (2b)$$

In the second-stage the predicted values of T_i are then used in equation (1) to instrument for whether or not a woman has more than two children.

6.3 Is same sex sibling composition a good instrumental variable in the context of South Africa?

In order for same sex sibling composition, S_i, to be a good instrument it must be i) uncorrelated with the error term, e_i, or exogenous within equation (1) and ii) correlated with

whether a woman has more than two births, T_i. If either of the two IV properties is not met then IV estimation can lead to large standard errors, asymptotic bias and inconsistent estimates (Wooldridge 2003; Bound *et al* 1995). For this reason I first test and discuss whether same sex sibling composition satisfies the two properties and is thus a good instrumental variable for further childbearing among African mothers in South Africa.

Previous studies have shown that same sex sibling composition satisfies the two properties of an instrumental variable in the United States, the United Kingdom, Argentina and Mexico (Angrist and Evans 1998; Iacovou 2001; Cruces and Galiani 2007). But same sex sibling composition may not be a good instrument for childbearing in South Africa. In particular, it is likely to fail the second IV property. Evidence from studies on sibling sex preferences does not suggest a preference for a mixed sibling sex composition in South Africa. In this case, no significant correlation may exist between same sex siblings and the endogenous explanatory variable, more than two births.

Using questions asked on desires for additional children in demographic health surveys from around the world, Arnold (1992) compares responses across women with same and mixed sibling sex composition. He argues that if preferences for a mixed sibling sex composition exist, then a smaller proportion of women with mixed sibling sex composition should want another child compared to women with a same sex sibling composition. However despite having already obtained a mixed sibling sex composition, the majority of women aged 15 to 49 in Sub-Saharan Africa wanted another child (Arnold 1992:94). By contrast, he finds evidence of a clear preference for mixed sibling sex composition among women with two children in Asian, Latin American and North African countries. A possible reason for the contrasting results is larger desired family sizes in Sub-Saharan African countries. Women may not be concerned about the sex of their first two children if they have desires for larger family sizes because eventually a mixed sibling sex composition will be obtained (Ben-Porath and Welch 1976:292). Arnold finds that desires for a balanced sex composition in Sub-Saharan African countries only begin to emerge among women with four consecutive children of the same sex (*ibid*:94).

Arnold's sample of Sub-Saharan African countries did not include South Africa. Gangadharan and Maitra (2003), however, conduct a specific study of sex preference in South Africa using the 1993 Project for Statistics on Living Standards and Development (PSLSD). They do not find a significant difference in the proportion of African mothers[36] in South Africa (with at least two children) who have a third birth, depending on whether the sex of their first two children was the same or mixed. Interestingly, they do find that White, Coloured and Indian women in South Africa, with at least two children, are more likely to have third birth if the first two children were of the same sex rather than mixed sex (*ibid*:388). A possible reason for this result is that population groups are heterogeneous in desired sex ratios and family size (Ben-Porath and Welch 1976). Coloureds, Indians and Whites have smaller family sizes than Africans, in which case they may be more concerned about the sibling sex composition of their first two children.

6.3.1 Property i) – exclusion restriction

It is generally difficult to directly test property i), otherwise called the exclusion restriction, because variables captured in the error term cannot be directly observed (Wooldridge 2003:496). However I use basic intuition as well as an analysis of mean contrasts between women with same sex and mixed sibling sex composition to validate the identification strategy.

There are concerns that in some developing countries the exclusion restriction may not hold due to sex-selection, for example through sex-selective abortion or stopping rules. [37] In this case the sex composition of children is non-randomly assigned and may possibly be correlated with background characteristics (Cruces and Galiani 2007:568; Clark 2000:95). Same sex sibling composition therefore may not be exogenous within a labour force

[36] In this study, Gangadharan and Maitra (2003:383) only identify mothers who are *co-resident* with their children using household relationship codes.

[37] Stopping rules, otherwise referred to as *differential stopping behaviour*, occur when couples stop bearing children when they have achieved the desired sex composition. In the case of son preference, couples will continue having children until they reach the desired number of sons (Clark 2000).

participation equation. Sex-selection is especially a concern in countries such as India and in North Africa in which strong preferences for sons are exhibited. Son preferences may be driven by tastes and lower net prices of sons compared to daughters (Ben-Porath and Welch 1976). The net cost of a son may be lower because he may be more likely to augment household income and provide old age support (Gangadharan and Maitra 2003:371). Dowries or bride prices may also raise the net cost of girls relative to boys (Clark 2000:95; Ben-Porath and Welch 1976:292). Among Africans in South Africa, however, bridewealth payments, called *ilobolo,* are traditionally made by the potential husband to the bride's family to validate a customary marriage (Casale and Posel 2007). In this case daughters become valuable to their parents for their associated *ilobolo* payments; and therefore the net cost of girls may not exceed the net cost of boys (Ganghadaran and Maitra 2003). For this reason a strong son preference may not exist among Africans in South Africa.

Accordingly, there is no empirical evidence to suggest a strong son preference among African parents in South Africa which could affect the biological random assignment of the sex composition of their offspring. Compared to high infant sex-ratios[38] in countries with strong son preferences, infant sex-ratios in South Africa are low and therefore not skewed in favour of boys (United Nations 2008). Using the 1996 Census, Garenne (2004) reports a sex-ratio of 0.991[39] for the African population under 12 years of age which is even lower than the established average of about 1.05 for European populations. Moreover if strong preferences existed for sons, then average birth intervals between successive births would tend to be shorter if the previous birth was a girl rather than a boy. Among Africans, however, Gangadharan and Maitra (2003) find no significant differences in birth intervals between successive births by the sex of the previous birth.[40] Furthermore, there is no evidence of discrimination against girls in terms of their schooling outcomes. Sibanda and

[38] Sex-ratios are calculated as the number of males to females.

[39] Consistent with the Census 1996, sex-ratios at birth using the Demographic Health Survey (DHS) 1998 and Agincourt data from 1992 to 2000 are reported at 0.995 and 0.997 respectively (Garenne 2004:95).

[40] However they do find evidence of shorter birth intervals after the birth of a daughter among Indians in South Africa suggesting a son preference exists among this group.

Lehloenya (2005), for example, find that gender differences in school enrolment rates are very small while primary school completion rates are in fact higher among girls than boys.

One way in which to check that the exclusion restriction holds is to compare the average demographic characteristics of women with a mixed sibling sex composition to women with same sex siblings (Aguero and Marks 2008; Angrist and Evans 1998). If there is no correlation between the exogenous instrument and the error term then there should be no systematic differences in the demographic characteristics between women whose first two children are of the same or mixed sex. Table 6.1 reports the mean differences in demographic characteristics for the two groups of women. Mean differences in age, age at first birth, marital status, rural/urban status and total household earnings do not differ from zero at the five percent level of significance. Specifically, labour force participation rates do not differ across the two groups of women.[41] However compared to co-resident mothers with a mixed sibling sex composition, a larger total number of children are born to co-resident mothers whose first two children born are of the same sex. Years of schooling are also slightly higher among co-resident mothers who have a mixed sibling sex composition compared to co-resident mothers with a same sex sibling composition, but only at the ten percent level of significance. Furthermore these educational attainment differences are observed only at lower levels of education.[42]

---

[41] Furthermore results of a basic regression of labour force participation on same sex sibling composition for a combined sample of co-resident and not co-resident African mothers (who have had at least two live births) show no statistically significant effect of sibling composition on labour force participation.

[42] It may be argued that women with higher levels of education are more likely to be aware of possible methods to influence the probability of conceiving a boy or a girl by planning intercourse at certain dates within an ovulation cycle. If women have a preference for a mixed sibling sex composition they may be able to exploit these methods to achieve the desired sex composition of children. For this reason the exclusion restriction would be violated because the likelihood of having a mixed sibling sex composition is influenced by background characteristics. But there are two problems with this reasoning. First, if women were to use these methods they would probably have higher levels of education such as a matric or postmatric education. But educational differences across the two groups of women are only observed at primary levels. Furthermore evidence that couples can influence the sex composition of children conceived by structuring intercourse at different points in the ovulation cycle remains inconclusive (Simpson 1995).

Table 6. 1: Differences in mean characteristics of African women by the sex composition of their first two births, 2002

Mean characteristic difference = mean (same sex siblings) – mean (mixed sex siblings)	Co-resident mothers having had at least two births	Not co-resident mothers having had at least two births
Age	-0.251 (0.171) t = -1.4664	0.693 (0.457) t = 1.516
Age at first birth	-0.066 (0.098) t = -0.6736	0.385 (0.251) t = 1.537
Married	0.003 (0.012) t = 0.2182	0.049 (0.032) t = 1.557
Rural	-0.004 (0.012) t = -0.2979	-0.015 (0.032) t = -0.471
Total monthly household earnings[a] (2002 prices)	5.538 (76.196) t = 0.0727	-36.383 (146.400) t = -0.2485
Years of schooling	0.195* (0.100) t = 1.9373	-0.120 (0.255) t = -0.469
No schooling	-0.017** (0.008) t = -2.2030	0.001 (0.020) t = 0.067
Incomplete Primary	-0.013 (0.010) t = -1.308	-0.026 (0.027) t = -0.982
Competed primary	0.014** (0.007) t = 2.006	-0.010 (0.020) t = -0.523
Incomplete Secondary	0.005 (0.012) t = 0.405	0.035 (0.031) t = 1.149
Matric	0.005 (0.009) t = 0.588	-0.007 (0.022) t = -0.328
Post-matric	0.006 (0.006) t = 0.883	0.007 (0.014) t = 0.496
Labour force participant	0.015 (0.012) t = 1.254	-0.030 (0.026) t = -1.174
No of births	-0.095*** (0.036) t = -2.640	0.0523 (0.083) t = 0.632

Source: GHS 2002

Notes: [a]Total monthly household earnings are in 2002 prices and include only earned income and not unearned income. The data are unweighted. Standard deviations are in parentheses. Means differences are calculated by subtracting the mean characteristic for women with mixed sex siblings from those with same sex siblings. T-statistic is abbreviated by 't'. The null hypothesis of the t-statistic is that the mean difference is zero i.e. no difference exists in the average characteristics of women with a same and mixed sibling sex composition. The sample includes African mothers aged 20 to 49 who have had at least two births. Co-resident mothers are living with at least one of their children aged 18 years or younger. Not co-resident mothers are not living with any of their children. *** 1 % significance level; ** 5% significance level; * 10% significance level.

Given the above discussion, there is little evidence that same sex sibling composition violates the exclusion restriction. There is no evidence to support a strong son preference which may interfere with the biological random assignment of the sex of children born to women. A comparison of mean characteristics across women with same sex and mixed sibling sex composition does not point to significant differences that would violate the exclusion restriction. Same sex sibling composition therefore is likely to satisfy property i) of an instrumental variable.

6.3.2 Property ii) – correlation between the instrumental variable and the endogenous explanatory variable

The mean contrasts in Table 6.1 indicated that among co-resident mothers with at least two children, more children are born to women whose first two children are of the same sex compared to women who have a boy and a girl. This suggests that there may in fact be a preference for a mixed sibling sex composition in South Africa. I test further whether the sibling sex composition of a mother's first two children influences fertility outcomes using descriptive estimates and first-stage estimation.

Table 6.2 compares the proportion of African women who have had more than two births by the sibling sex composition of their first two children. First, I consider results for a combined or aggregate sample of all mothers, regardless of residency with children and marital status (see first row division in Table 6.2). In 2002, the proportion of African women having had a third birth was statistically higher among mothers whose first two children were of the same sex rather than different sex. This contrasts with the findings of Gangadharan and Maitra (2003) using the PSLSD 1993.

Table 6. 2: Proportion of mothers who have had a third birth, African mothers 2002

	All mothers (not distinguished by co-residency with children)		
Sex of first two children among women with two or more children	**Married and unmarried**	**Married**	**Unmarried**
(1) Mixed sibling sex	0.602 (0.490) N = 2,223	0.682 (0.466) N = 1,497	0.484 (0.500) N = 726
(2) Same sex sibling	0.620 (0.485) N = 2,363	0.697 (0.460) N = 1,547	0.513 (0.500) N = 816
Two girls	0.623 (0.485) N = 1,186	0.698 (0.459) N = 797	0.525 (0.500) N = 389
Two boys	0.622 (0.485) N = 1,191	0.702 (0.457) N = 763	0.504 (0.500) N = 428
Mean difference (2) – (1)	0.019** (0.011) t = 1.6530	0.016 (0.014) t = 1.1132	0.029* (0.018) t = 1.601
	Co-resident mothers		
Sex of first two children among women with two or more children	**Married and unmarried**	**Married**	**Unmarried**
(3) Mixed sibling sex composition	0.617 (0.486) N = 1,990	0.694 (0.461) N = 1.369	0.496 (0.500) N = 621
(4) Same sex sibling composition	0.638 (0.481) N = 2,099	0.710 (0.453) N = 1,421	0.526 (0.500) N = 678
Two boys	0.635 (0.481) N = 1,048	0.711 (0.454) N = 728	0.512 (0.500) N = 320
Two girls	0.646 (0.478) N = 1,065	0.716 (0.451) N = 706	0.543 (0.500) N = 359
Mean Difference (4) – (3)	0.021** (0.012) t = 1.7373	0.017 (0.015) t = 1.1525	0.030* (0.020) t = 1.4933
	Not co-resident mothers		
Sex of first two children among women with two or more children	**Married and unmarried**	**Married**	**Unmarried**
(5) Mixed sibling sex composition	0.496 (0.501) N = 233	0.577 (0.495) N = 128	0.424 (0.495) N = 105
(6) Same sex sibling composition	0.510 (0.500) N = 264	0.580 (0.495) N = 126	0.458 (0.499) N = 138
Two boys	0.535 (0.500) N = 126	0.619 (0.488) N = 57	0.469 (0.501) N = 69
Two girls	0.485 (0.501) N = 138	0.538 (0.501) N = 69	0.448 (0.499) N = 69
Mean Difference (6) – (5)	0.014 (0.032) t = 0.4354	0.003 (0.047) t = 0.0707	0.034 (0.043) T = 0.8100

Source: GHS 2002

Notes: Data are unweighted. Standard deviations are in parentheses. Sample is of African mothers aged 20 to 49 who have had *at least two births*. A co-resident mother is a woman co-residing with at least one of her own children aged 18 years or younger. A not co-resident mother is a woman with own children aged 18 years or younger but not living with any of her children. The alternative hypothesis is that the mean difference is greater than zero. *** 1 % significance level; ** 5% significance level; * 10% significance level.

However preferences for sibling sex composition differ by marital status and whether mothers are co-resident or not co-resident with their children. The 'aggregate' preference among mothers for a mixed sibling sex composition is driven specifically by the sample of co-resident mothers who are unmarried. Among this group, who constitute 34 percent of the total sample of African mothers with at least two births, the proportion having had a third birth is statistically higher among mothers whose first two children were of the same sex rather than opposite sex. However, among married co-resident mothers and all mothers not co-resident with their children, there are no significant mean contrasts in the proportion having had a third birth by sibling sex composition. I however acknowledge that insignificant mean contrasts among the sample of not co-resident mothers may be due to small sample sizes.

In addition to descriptive statistics, first-stage estimations can be used to directly test property ii) – whether a correlation exists between the endogenous explanatory variable and the instrument. As specified in equations (2a) and (2b), I estimate the relationship between more than two births and sibling sex composition. The results are presented in Table 6.3.

Even after partialling out the effects of other explanatory variables included in the reduced form equation, the first-stage results in Table 6.3 are generally consistent with the descriptive results in Table 6.2. Considering the combined sample of all mothers (not distinguished by residency with children and marital status) the aggregate coefficient on same sex siblings is positive and significant at the five percent level but small in magnitude. Compared to mothers with a mixed sibling sex composition, mothers with a same sex sibling composition are only about two percent more likely to have a third birth. The results also indicate that on average there is a bias for boys: the coefficient on two girls is larger and more significant than the coefficient on two boys. Therefore compared to mothers whose first two children are boys, mothers whose first two children are girls are more likely to have a third birth in the hope of conceiving a boy.

Table 6. 3: First-stage results - dependent variable is *more than two births*, T_i

	All mothers (not distinguished by co-residency with children)		
	Married and unmarried	**Married**	**Unmarried**
Instrument: Same sex siblings			
Coefficient – same sex	0.020** (0.009)	0.011 (0.012)	0.031** (0.015)
Partial R^2	0.0006	0.0002	0.0014
F-statistic	4.541	0.822	4.216
Instrument: Two girls, two boys			
Coefficient – Two Girls	0.023** (0.011)	0.010 (0.014)	0.043** (0.018)
Coefficient – Two Boys	0.016 (0.011)	0.011 (0.014)	0.019 (0.019)
Partial R^2	0.0006	0.0002	0.0018
F-statistic	2.406	0.384	2.751
N	7,477	4,389	3,088
	Co-resident mothers		
	Married and unmarried	**Married**	**Unmarried**
Instrument: Same sex siblings			
Coefficient – same sex	0.018* (0.010)	0.011 (0.012)	0.026* (0.017)
Partial R^2	0.0005	0.0002	0.0010
F-statistic	3.24202	0.840	2.499
Instrument: Two girls, two boys			
Coefficient – Two Girls	0.026** (0.012)	0.013 (0.015)	0.045** (0.020)
Coefficient – Two Boys	0.009 (0.012)	0.008 (0.015)	0.008 (0.021)
Partial R^2	0.0007	0.0002	0.0020
F-statistic	2.303	0.419	2.508
N	6,487	3,949	2,538
	Not co-resident mothers		
	Married and unmarried	**Married**	**Unmarried**
Instrument: Same sex siblings			
Coefficient – same sex	0.039 (0.028)	0.015 (0.041)	0.056 (0.037)
Partial R^2	0.0021	0.0003	0.0044
F-statistic	2.019	0.128	2.30815
Instrument: Two girls, two boys			
Coefficient – Two Girls	0.017 (0.034)	-0.012 (0.051)	0.048 (0.045)
Coefficient – Two Boys	0.062* (0.034)	0.041 (0.051)	0.065 (0.045)
Partial R^2	0.0035	0.0022	0.0046
F-statistic	1.674	0.452	1.203
N	990	440	550

Source: GHS 2002

Notes: Data are unweighted. Standard deviations are in parentheses. Sample is of African mothers aged 20 to 49 who have had *at least two births*. A co-resident mother is a woman co-residing with at least one of her own children aged 18 years or younger. A not co-resident mother is a woman with own children aged 18 years or younger but not living with any of her children. Other exogenous variables include the woman's age, age squared, age at first birth, her educational status and marital status. Controls are also included for household compositional variables, dummy variables for province of residence and whether the woman lives in a rural or urban area. *** 1 % significance level; ** 5% significance level; * 10% significance level.

Again the aggregate results are driven *only* by the sample of *unmarried* mothers who are co-resident with at least one of their children. The coefficients on sibling sex composition variables are almost always insignificant for the sample of married[43] co-resident mothers and for both married and unmarried not co-resident mothers. Therefore heterogenous preferences exist for desired sibling sex composition among the different samples of mothers. This may be attributed to differences in family size across the different groups. Table 6.4 shows that compared to the group of married co-resident mothers, unmarried co-resident mothers have on average fewer children and may therefore be more concerned about the sex composition of their first few children.

An interesting result from the first-stage estimations is the bias for boys among unmarried mothers who are co-resident with at least one of their own children. This bias for boys emerges despite a context of bridewealth payments which raises the value of daughters. A possible explanation is that in the absence of a spouse/partner to provide financial support, unmarried co-resident mothers may have a preference for sons who may be better positioned in the market place than daughters to support their mothers in old age. But no son preference is evident among the sample of mothers who are unmarried but are not co-residing with any of their children. One reason for this result is that unmarried mothers who are not co-resident with their children may be more likely to be working and therefore more economically independent. This may mitigate the need for financial support from a son. Table 6.4 indicates that 84 percent of unmarried not co-resident mothers (with at least two children) were labour force participants compared to only 69 percent of unmarried co-resident mothers (with at least two children).

[43] It may be criticised that statistically insignificant coefficients are obtained on sibling sex variables for the sample of married mothers because included in the sample are married mothers who are not living with their spouse or partner. Married parents who do not live together may be less likely to form preferences for family size and composition than married parents who live apart. But I find that results are robust to the inclusion or exclusion of married mothers who are not living with their spouse or partner in the sample of married women.

Table 6. 4: Number of live births and labour force participation rates among African mothers having had at least two births, 2002

	Co-resident mothers		Not co-resident mother	
	Married	**Unmarried**	**Married**	**Unmarried**
Number of births	3.631	2.986	3.259	2.803
	(1.559)	(1.275)	(1.437)	(1.136)
Labour force participation	0.594	0.685	0.736	0.837
	(0.491)	(0.464)	(0.441)	(0.370)
N	3,949	2,538	440	550

Source: GHS 2002

Notes: Data are unweighted. Standard deviations are in parentheses. Sample is of African mothers aged 20 to 49 who have had *at least two births*. A co-resident mother is a woman co-residing with at least one of her own children aged 18 years or younger. A not co-resident mother is a woman with own children aged 18 years or younger but not living with any of her children.

Despite a statistically significant relationship between same sibling sex composition and having a third birth among the combined or aggregate sample of mothers, this result is driven only by a select sample of mothers. Sibling sex composition is a poor instrument for further childbearing among married co-resident mothers and all not co-resident mothers. Furthermore, the quality of the instrumental variable is poor given the small values of the partial R^2 and F-statistic of the identifying instruments in the first-stage estimation (Bound *et al* 1995). This result holds regardless of the sample used. The partial R^2 values measure the strength of the linear relationship between the endogenous explanatory variable and the identifying instrument once the effects of other exogenous variables have been partialled out (Wooldridge 2003:490, Bound *et al* 1995:444). The partial R^2 value of the identifying instruments, same sex sibling composition or two girls and two boys, is less than 0.005 for every sample identified in Table 6.3. Moreover, F-statistics for the single instrument, same sex siblings, always lie well below the suggested value of ten for a strong instrument (Stock *et al* 2002:522). For this reason I conclude that sibling sex composition is a weak instrument for further childbearing among mothers in South Africa, regardless of co-residency with children or marital status.

As a consequence of weak instruments, the IV estimates of the effects of childbearing on labour force participation are likely to be more inconsistent than OLS estimates and yield imprecise coefficients with no meaningful interpretation. I illustrate this in the next section.

6.3.2 Ordinary least squares and second-stage results

In Table 6.5 I present second-stage IV estimates using the single variable, same sex sibling composition, or two variables, two girls and two boys, as instruments for a third birth. I also present OLS estimates for comparative purposes. The table splits the sample according to residency with children and marital status. Considering the aggregate sample of mothers (not distinguished by residency with children and marital status) OLS results show that having a third birth does not have a statistically significant effect on a mother's labour force participation. When disaggregated by marital status, the OLS estimate on *more than two births* is significant but differs in sign depending on marital status. Married mothers are significantly less likely to be labour force participants if they have a third birth while unmarried mothers are significantly more likely to be labour force participants if they have a third birth. However the results for the combined sample of co-resident and not co-resident mothers are driven only by the sample of co-resident mothers. There is no effect of having a third birth on the labour force participation of not co-resident mothers.

IV estimates differ considerably from OLS estimates; they are 'very noisy' and are estimated with little precision. Moreover IV estimates provide evidence of implausible coefficients. For example, IV estimate results using same sex siblings in the second row division of Table 6.5 indicate that among co-resident mothers who are unmarried, having a third child *lowers* their labour force participation by as much as 93 percent. However this estimate lies within a 95 percent confidence interval ranging from negative 268 percent to a positive value of 83 percent. The IV estimate therefore cannot be interpreted with any level of confidence. Among the same group of women, the OLS estimate indicates that they are five percent *more* likely to be labour force participants if they have a third child. This OLS estimate can be interpreted with more precision than the IV estimate because its 95 percent confidence interval is considerably smaller, ranging from one percent to nine percent (confidence intervals are not presented in the table).

The noisy and imprecise IV estimates in Table 6.5 can be attributed to the weak explanatory power of sibling sex composition variables in the first-stage estimations; and

the cost of this weak explanatory power is very large IV standard errors in the second-stage. Standard errors of IV estimates are always expected to be greater than those of OLS estimates due to imperfect correlation between the endogenous explanatory variable and the identifying instrument (Wooldridge 2003). Consider the equation below for the asymptotic variance (or square of the asymptotic standard error) of an IV estimate,

$\sigma^2 / (SST_T * R^2)$

where SST_T is the total sum of squares of T_i, σ^2 is the variance of e_i and R^2 measures the strength of the linear relationship between the having a third child and the identifying instrument, same sex sibling composition. An R^2 value of less than one will always raise the IV standard error above that of the OLS standard error (Wooldridge 2003:490). First-stage estimation results showed that partial R^2 values were always less than 0.005 (see Table 6.3). Given these very small R^2 values, it is not surprising that IV standard errors are 20 to 200 times greater than the standard errors of OLS estimates. For example, among not co-resident mothers who are married, the standard error on the OLS estimate is 0.05 while the standard error on the IV estimate using same sex siblings is 200 times larger at 10.23 (see third row division in Table 6.5).

As a result of large standard errors, IV estimates cannot be interpreted with any level of confidence and are meaningless (Wooldridge 2003:494). Bound *et al* (1995:443) note that even if the estimations could be run with larger sample sizes, reducing standard errors of IV estimates, the weak correlation between endogenous explanatory variable and instrument would result in IV estimates that are biased in the same direction as OLS estimates (Bound *et al* 1995:443).

Table 6. 5: Second-stage and OLS results - dependent variable is *labour force participation*, Y_i

	All mothers (not distinguished by co-residency with children)		
	Married and unmarried	**Married**	**Unmarried**
OLS	-0.013 (0.013)	-0.051*** (0.018)	0.037** (0.018)
IV: same sex	-0.230 (0.525)	0.599 (1.481)	-0.546 (0.558)
IV: two girls and two boys	-0.038 (0.501)	0.346 (1.415)	-0.083 (0.424)
N	7,477	4,389	3,088
	Co-resident mothers		
	Married and unmarried	**Married**	**Unmarried**
OLS	-0.007 (0.014)	- 0.048** (0.019)	0.048** (0.021)
IV: same sex	-0.565 (0.700)	0.099 (1.316)	-0.928 (0.895)
IV: two girls and two boys	0.032 (0.527)	0.145 (1.324)	-0.023 (0.459)
N	6,487	3,949	2,538
	Not co-resident mothers		
	Married and unmarried	**Married**	**Unmarried**
OLS	0.016 (0.028)	-0.020 (0.050)	0.030 (0.034)
IV: same sex	0.631 (0.757)	3.601 (10.233)	-0.003 (0.500)
IV: two girls and two boys	0.153 (0.487)	0.306 (1.120)	-0.080 (0.493)
N	990	440	550

Source: GHS 2002

Notes: Data are unweighted. Standard errors are in parentheses. Sample is of African mothers aged 20 to 49 who have had *at least two births*. A co-resident mother is a woman co-residing with at least one of her own children aged 18 years or younger. A not co-resident mother is a woman with own children aged 18 years or younger but not living with any of her children. Other exogenous variables include the woman's age, age squared, age at first birth, her educational status and marital status. Controls are also included for household compositional variables, dummy variables for province of residence and whether the woman lives in a rural or urban area. *** 1 % significance level; ** 5% significance level; * 10% significance level.

Conclusion

International studies have argued that same sex sibling composition is a good instrument for further childbearing among mothers with at least two children in the United States, the United Kingdom, Argentina and Mexico. In these studies same sex sibling composition satisfies the two properties of an instrumental variable (Cruces and Galiani 2007; Iacovou

2001; Angrist and Evans 1998). Consistent with these studies there is little evidence to show that sibling sex composition among Africans in South Africa violates property i) (or the exclusion restriction). On aggregate, there is no evidence of strong sex-selection in South Africa which would affect the random assignment of a child's sex. Furthermore, a comparison of characteristics across mothers with at least two children, by the sibling sex composition of their first two births, does not provide evidence that sibling sex composition is influenced by background characteristics which may violate the exclusion restriction.

However in contrast to international studies, using sibling sex composition to instrument for further childbearing among Africans in South Africa does not adequately satisfy property ii) – whether there is a strong correlation between the endogenous explanatory variable and the instrument. Same sex sibling composition has weak explanatory power in first-stage estimations of the propensity to have more than two children. Partial R^2 values and F-statistics are very low in first-stage estimations, regardless of residency with children or marital status. Furthermore any correlations observed between sibling sex composition variables and having more than two children for the aggregate sample of mothers is driven only by a select sample of unmarried mothers who are co-resident with their children.

Same sex sibling composition is therefore a weak instrument for childbearing among African women in South Africa. A possible reason for this result is that populations are heterogeneous in desired sex-ratios and family size (Ben-Porath and Welch 1976:292). In particular family sizes among Africans in South Africa are generally larger than in the other countries in which same sex sibling composition has been used to instrument for childbearing (Central Intelligence Agency 2008). With larger desired family sizes, African women may not be particularly concerned about the sex of their first two children because they expect to achieve a mixed sibling sex composition when more children are born (Ben-Porath and Welch 1976). This hypothesis may also explain differences in observed correlations between further childbearing and sibling sex composition across the different samples of African mothers, disaggregated by co-residency with children and marital status.

The consequence of a weak instrumental variable is that IV estimates of the effect of further childbearing on women's labour force participation are more inconsistent than OLS estimates. The weak explanatory power of the sibling sex composition instruments in first-stage estimations translates into very large standard errors of IV estimates in the second-stage. In this analysis, standard errors on IV estimates were 20 to 200 times larger than standard errors of OLS estimates. The resulting IV estimates are therefore 'very noisy', are estimated with little precision and have no meaningful interpretation.

Even if sibling sex composition was a good instrument for childbearing, it would only be able to estimate the causal effect of having a third birth on the labour force participation of women *with at least two children.* But the effect of a third birth on women's labour force participation may be far smaller than the effect of the first child. A better identification strategy to identify the causal relationship between having children and labour force participation would use infertility status as an instrument for childbearing. First suggested by Aguero and Marks (2008), infertility status allows the researcher to determine the effects of the first, second and proceeding births on labour force participation. If data on infertility status were made available in South African household surveys, it would enable further research on the causal relationship between motherhood and women's labour force participation in South Africa.

Conclusion

The primary objective of this thesis was to explore the relationship between motherhood and female labour force participation in South Africa. Although a vast international literature exists on this relationship, the effects of motherhood had not been analysed in South Africa. The reason for the dearth of literature on this topic is that datasets used in studies of female labour force participation in South Africa have not contained information to match women to their children. A secondary (but obviously related) objective of the thesis was to analyse what datasets are available in South Africa and how they could be used to match women to their children. I outline the key findings and contributions related to the primary thesis objective followed by those relating to the secondary thesis objective.

International studies have typically placed a co-residency restriction on motherhood when estimating the relationship between motherhood and female labour force participation. Using a sample of African women aged 20 to 49 in the General Household Survey 2002 I showed that if mothers are defined as women who are *co-resident* with at least one of their biological children, there is evidence that motherhood has a significant negative effect on whether women participate in the South African labour force. Furthermore this negative relationship is driven particularly by the presence of very young children.

Using ordinary least squares (OLS) estimation, I find a negative motherhood effect of six percent when a co-residency restriction is placed on the definition of motherhood. This effect is comparatively smaller than observed motherhood effects in other countries when studies do not control for the endogeneity of motherhood in labour force estimations (Aguero and Marks 2008; Boushey 2008). It was highlighted that the size of this negative motherhood effect in South Africa may be biased for two reasons. First, the coefficient on motherhood variables may also capture the indirect effects of omitted variables in labour force equations. For example, the omission of ambition or motivation, which are positively correlated with labour force participation but negatively correlated with motherhood status, means that the negative motherhood effect may be over-estimated. Second, a co-residency restriction on the definition of motherhood may overestimate the size of the negative

motherhood effect if co-resident mothers are a non-random sample of all mothers who are less likely to participate in the labour force.

This thesis provided clear evidence that the definition of motherhood makes a particularly big difference to the estimated motherhood effect. A significant sample of all African mothers are not co-resident with any of their children: in 2002, about 15 percent of all African mothers aged 20 to 49 were living apart from all their children aged 18 years or younger. A key reason for the absence of mothers is labour migration to places of employment where women leave children in the household of origin. The study showed that these absent or not co-resident mothers are significantly more likely than both co-resident mothers and non-mothers to be labour force participants. The remaining sample of co-resident mothers, who are more likely to be married and living in rural areas, are therefore a non-random sample of all mothers. Restricting the definition of motherhood to only co-resident mothers biases upward the negative relationship between motherhood and women's labour force participation. In fact when the co-residency restriction on motherhood is removed, the negative motherhood effect falls from six percent to about two percent and is only weakly significant.

Despite estimating a small 'aggregate' motherhood effect of only negative two percent, the size of this effect may still be biased in the presence of unobserved heterogeneity. In accordance with international literature, I considered an instrumental variable approach to address this possible source of endogeneity bias. Specifically I followed the methodology of Angrist and Evans (1998) who exploit the sibling composition of a woman's first two children to identify the exogenous effects of childbearing on labour force participation. They suggest that individuals have a preference for a mixed sibling sex composition which induces them to have a third birth if their first two offspring are of the same sex (Ben-Porath and Welch 1976). Therefore sibling sex composition is correlated with a woman's childbearing behaviour and would be expected to provide a good instrument for further childbearing. In particular sibling sex composition has been found to be a strong instrument for childbearing in the United States (Angrist and Evans 1998), the United Kingdom (Iacovou 2001), Mexico and Argentina (Cruces and Galiani 2007). However, contrary to

international studies, this thesis showed that sibling sex composition is a poor instrument for further childbearing among Africans in South Africa. First-stage results of the two-stage least squares estimation exhibited a weak correlation between same sex sibling composition and the endogenous explanatory variable. A possible explanation for this result is that compared to average family sizes in the aforementioned countries, Africans in South Africa have larger families. Women may not be concerned about the sex of their first two children as they eventually expect to achieve a mixed sibling sex composition.

The weak correlation between same sex sibling composition and further childbearing translated into very large standard errors on IV estimates of the effects of childbearing. Consequently, IV estimates were found to be very 'noisy', they lacked precision and had no meaningful interpretation. For this reason, using same sex sibling composition as an instrumental variable fails to identify the exogenous effect of childbearing on labour force participation in South Africa. Compared with IV estimates, the 'aggregate' negative motherhood effect of two percent, estimated using OLS, may provide a more reliable estimate of the motherhood effect in South Africa.

In addition to exploring the relationship between childbearing and female labour force participation, a related objective of this thesis was to investigate whether the provision of surrogate parenthood or foster care constrained women's labour force participation. A key characteristic of the African sample of women used in this study, is that a significant proportion of these women live with children aged 18 years or younger who are not their own. Furthermore many of the children with whom they reside have absent mothers. Two thirds of all African non-mothers aged 20 to 49 were living in a household with other children aged 18 years or younger in 2002. And among this group of non-mothers specifically, 46 percent were living in a household with at least one child whose mother was absent from the household. Furthermore among co-resident mothers aged 20 to 49, 46 percent were also living with children who were not their own in 2002. These findings mirror reports of extended family living arrangements and child fosterage among African households.

Chapter Five indicated that living with other children does not strengthen the negative relationship between motherhood and labour force participation among co-resident mothers. However, non-mothers who live with children are significantly less likely to be labour force participants than non-mothers who are not living with any children. Although the size of the child effect is smaller for non-mothers than for co-resident mothers, the results indicate that the costs of childcare are not borne by mothers alone.

A key challenge in this thesis was to identify among women who were mothers using household surveys. This was a first step to investigating the effects of childbearing (and surrogate parenthood) on women's labour force participation. In previous studies of female labour participation in South Africa, the effects of childbearing could not be analysed because datasets used could not match women to their children. In direct response to this problem, the thesis analysed what data are available to match women to their children and the quality of these data. A contribution of this analysis was to outline four possible methods of identifying mothers using either birth status information or questions on household relationships, highlighting their benefits and limitations. The international literature has typically used a household relationship question to identify mothers. However, a sizeable sample of mothers is not co-resident with their children in South Africa. Furthermore, there has been an increase in the proportion of mothers living apart from all their children between 1993 and 2002. This is consistent with a reported rise in female labour migration over this period (Posel and Casale 2006; Collinson *et al* 2003). Because a large (and increasingly large) sample of mothers is not co-resident with their children, household relationship information is inadequate to identify the extent of motherhood in South Africa. This is particularly the case among African women who are more likely than women in other race groups to be living apart from their children.

Although birth status information is most effective in identifying mothers using South African household surveys, the thesis illustrated that the quality of birth information in cross-sectional and panel surveys is compromised by high levels of non-response, especially among adolescents. Future studies on motherhood and its economic costs in South Africa will require surveys containing reliable birth status information to identify

representative samples of mothers. For this reason effort should be taken to design and administer birth module questions which minimise problems of non-response and reporting error. An El-Badry (1961) adjustment to Census 1996 data provided evidence that a pervasive problem in collecting birth data in South Africa are 'zero-errors' in reporting where enumerators incorrectly code non-mothers as non-respondents. These errors could be prevented through training enumerators to fill out accurately answers to birth module questions.

However, improving the quality of birth data in panel surveys is further compromised by measurement error and sample attrition over repeated surveys. This was illustrated using the first two waves of the KwaZulu-Natal Income Dynamics Study in which about forty percent of women in a study of motherhood fell out of the panel between 1993 and 1998. Furthermore, there was evidence that sample attrition was selective on observed characteristics, including a woman's fertility/childbearing status. A key reason for high levels of sample attrition was that resident household members in 1993 who subsequently became non-residents in 1998 were not questioned on their birth status. Future panel surveys which collect birth data should attempt to follow and interview women who become non-resident household members. This may be costly but will considerably reduce sample attrition and prevent sample selection on observable characteristics.

The availability and quality of birth data in panel surveys will be particularly important for future studies of motherhood and its economic costs in South Africa. An area for further research would be to identify the opportunity cost of motherhood in terms of forgone employment hours and earnings. International studies that identify an employment and wage penalty to motherhood typically rely on panel data to identify unbiased estimates of the effects of motherhood (see for example Anderson et al 2003, 2002; Joshi 2002; Budig and England 2001; Waldfogel 1997). As identified in this thesis, unobserved heterogeneity complicates estimations of the opportunity costs of motherhood. Unobserved characteristics that may be correlated with a woman's motherhood status may simultaneously be correlated with her employment or earnings potential. Where unobserved characteristics remain constant over time they can be differenced out of an equation using fixed effects

panel models. Controlling for unobserved heterogeneity in this manner, it is possible to identify an exogenous effect of motherhood on a woman's employment or earnings potential.

The forthcoming National Income Dynamics Study (NIDS), which is the first nationally representative panel study in South Africa, will provide the necessary data to estimate the opportunity costs of motherhood in South Africa. NIDS will simultaneously identify birth information and a large variety of socio-economic variables for 8,000 households. Furthermore a key feature of the panel study is its ability to follow people as they move out of the original 8 000 households (National Income Dynamics Study 2008); therefore it will attempt to track migrants and specifically mothers who move away from their children in the household of origin. In particular, it could be used to extend this research on the relationship between motherhood and women's labour force participation. In the thesis I examined this relationship in a static context; but motherhood is a dynamic process where its effect on labour force participation is expected to change over the course of the life cycle. Panel data could be used to analyse this relationship in a dynamic context. For example, it would be possible to identify how transitions into motherhood and the birth of additional children may influence the labour force participation of women over time.

NIDS may also make available data that could be used as better instruments for childbearing compared with sibling sex composition. The study showed that sibling sex composition is a weak instrument for childbearing among Africans; furthermore it cannot be used to instrument for a woman's first or second birth i.e. it cannot identify a mother from a non-mother. An instrument for childbearing that is highly correlated with motherhood but unrelated to labour force participation is women's infertility status. This instrument is able to identify if a woman has had at least one birth. If data on infertility status are made available during the successive waves of the NIDS panel, it may allow researchers to identify more accurately the exogenous effects of motherhood on women's labour force participation.

References

Aguero, J. and Marks, M. (2008) "Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks". *American Economic Review* 98(2):500-504.

Alderman, H., Behrman, J., Kohler, H., Maluccio, J. and Watkins, S. (2001) "Attrition in Longitudinal Household Survey Data". *Demographic Research* 5:79-124.

Anderson, B. and Phillips, H. (2006) "Trends in the percentage of children who are orphaned in South Africa: 1995–2005". Pretoria: Statistics South Africa, Report No. 03-09-06.

Anderson, D., Binder, M. and Krause, K. (2002) "The Motherhood Wage Penalty: Which Mothers Pay and Why?" *The American Economic Review* 92(2):345-358.

Anderson, D., Binder, M. and Krause, K. (2003) "The Motherhood Wage Penalty Revisited: Experience, Heterogeneity, Work Effort and Work-Schedule Flexibility". *Industrial and Labour Relations Review* 56(2):273-294.

Angrist, J. and Evans, W. (1998) "Children and their parents' labour supply: Evidence from exogenous variation in family size". *The American Economic Review* 88(3):450-477.

Angrist, J. and Krueger, A. (2001) "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments". *The Journal of Economic Perspectives* 15(4):69-85.

Amoateng, A., Heaton, T. and Kalule-Sabiti, I. (2007) "Living arrangements in South Africa". In: Amoateng, A. Y. and Heaton, T. B. (eds) *Families and households in post-apartheid South Africa.* Cape Town: HSRC Press, pp. 43 – 59.

Ardington, C., Case, A. and Hosegood, V. (2007) "Labour Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa". National Bureau of Economic

Research Working Paper Series 13442. (Accessed March 2008) Available: http://www.nber.org/papers/w13442

Arnold, F. (1992) "Sex Preferences and Its Demographic and Health Implications". *International Family Planning Perspectives* 18(3):93-101.

Becker, G. (1960) "An Economic Analysis of Fertility". In: Universities National Bureau Committee for Economic Research (ed.) *Demographic and Economic Change in Developed Countries.*

Becker, G. (1965) "A Theory of the Allocation of Time". *The Economic Journal* 75(299): 493-517.

Becker, G. (1985) "Human Capital, Effort, and the Sexual Division of Labour". *Journal of Labour Economics* 3(1, Part 2):S33-S55.

Ben-Porath, Y. and Welch, F. (1976) "Do sex preferences really matter?" *The Quarterly Journal of Economics* 90(2):285-307.

Bhorat, H. and Leibbrandt, M. (2001) "Modelling Vulnerability and Low Earnings in the South African Labour Market". In: Bhorat, H., Leibbrandt, M., Maziya, M., van der Berg, S and Woolard, I. *Fighting Poverty. Labour market and Inequality in South Africa.* Cape Town: UCT Press. pp. 107 – 129.

Bianchi, S. (2000) "Maternal employment and time with children: Dramatic change or surprising continuity?" *Demography* 37(4):401 – 414.

Blacker, J. (1977) "The estimation of adult mortality in Africa from data on orphanhood". *Population Studies* 31:107-128.

Bongarts, J. (2003) "Completing the Fertility Transition in the Developing World: The Role of Educational Difference and Fertility Preferences". *Population Studies* 57(3):321-335.

Bound, J., Jaeger, D. and Baker, R. (1995) "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak". *Journal of the American Statistical Association* 90(430):443-450.

Boushey, H. (2008) "'Opting Out?' The Effect of Children on Women's Employment in the United States". *Feminist Economics* 14(1):1 – 36.

Bray, R. and Brandt, R. (2007) "Child Care and Poverty in South Africa". *Journal of Children and Poverty* 13(1):1-19.

Browning, M. (1992) "Children and Household Economic Behaviour". *Journal of Economic Literature* 30(3):1434-1475.

Budig, M. and England, P. (2001) "The Wage Penalty for Motherhood". *American Sociological Review* 66(2):204-225.

Brewster, K. and Rindfuss, R. (2000) "Fertility and women's employment in industrialized nations". *Annual Review of Sociology* 26:271-296.

Bronars, G. and Grogger, J. (1994) The economic consequences of unwed motherhood: Using twin births as a natural experiment. *The American Economic Review* 84(5):1141 – 1156.

Cahuc, P. and Zylberberg, A. (2004) *Labour Economics*. Cambridge, Massachusetts: The MIT Press.

Cain, G. and Dooley, M. (1976) "Estimation of a Model of Female Labour Supply, Fertility, and Wages of Married Women. *Journal of Political Economy*" 84(4):S179-S199.

Casale, D. (2003) "The Rise in Female Labour Force Participation in South Africa: An analysis of Household Survey Data, 1995 – 2001". PhD Thesis at the University of Natal, Durban.

Casale, D. and Posel, D. (2003) "The Continued Feminisation of the Labour Force in South Africa: An analysis of Recent Data and Trends". *The South African Journal of Economics* 70(1):156-183.

Casale, D. (2004) "What has the Feminisation of the Labour Market 'Bought' Women in South Africa? Trends in Labour Force Participation, Employment and Earnings, 1995-2001". *The Journal of Interdisciplinary Economics* 15:251-275.

Casale, D. and Posel, D. (2007) "The Male Marital Earnings Premium in the Context of Bridewealth Payments: Evidence from South Africa". ERSA Working Paper, no. 57.

Chapman, B., Dunlop, Y., Gray, M., Liu, A. and Mitchell, D. (2001) "The impact of children on the earnings lifetime of Australian women: Evidence from the 1990's". *The Australian Economic Review* 34(4): 373 – 389.

Central Intelligence Agency (2008) "World Fact Book: Rank Order – Total Fertility Rate". (Accessed 10 September 2008). Available: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2127rank.html

Clark, S. (2000) "Son Preference and Sex Composition of Children: Evidence from India". *Demography* 37(1):95-108.

Collinson, M., Tollman, S., Kahn, K. and Clark, S. (2003) "Highly prevalent circular migration: Households, mobility and economic status in rural South Africa". Paper

prepared for Conference on African Migration in Comparative Perspective, Johannesburg, South Africa, 4-7 June, 2003.

Collinson, M. and Wittenberg, M. (2001) "Labour Force Dynamics in a Rural Part of South Africa: the Agincourt sub-district of the Northern Province, 1992-2000". Presented at the DPRU/FES Conference, "Labour Markets and Poverty in South Africa", Johannesburg.

Connelly, R. (1992) "The Effect of Child Care Costs on Married Women's Labour Force Participation". *The Review of Economics and Statistics* 74(1):83-90.

Connelly, R., DeGraff, D., Levinson, D. and McCall, B. (2006) "Tackling the Endogeneity of Fertility in the Study of Women's Employment in Developing Countries: Alternative Estimation Strategies Using Data From Urban Brazil". *Feminist Economics* 12(4):561-597.

Cramer, J. (1979) "Employment Trends of Young Mothers and the Opportunity Cost of Babies in the United States". *Demography* 16(2):177-197

Cruces, G. and Galiani, S. (2007) "Fertility and female labor supply in Latin America: New causal evidence". *Labour Economics* 14(3):565 – 573.

Del Boca, D. and Locatelli, M. (2006) "The Determinants of Motherhood and Work Status: A Survey". Institute for the Study of Labour (IZA) Discussion Paper Series No. 2414.

Denis, P and Ntsimane, R. (2006) "Chapter 19: Absent fathers: why do men not feature in stories of families affected by HIV/AIDS in KwaZulu-Natal?" In: Richter, L. and Morrell, R. *Baba: Men and Fatherhood in South Africa.* Cape Town: HSRC Press. pp. 237 – 249.

Desai, S. and Jain, D. (1994) "Maternal employment and changes in family dynamics: The social context of women's work in rural South India". *Population and Development Review* 20(1):115 – 136.

Deustch, R. (1998) "Does Child Care Pay? Labour Force Participation and Earnings Effects of Access to Cild Care in the Favelas of Rio de Janiero". Inter-American Development Bank Working Paper no. 384.

Doepke, M., Hazan, M. and Maoz, Y. (2007) "Baby Boom and World War II: a macroeconomic analysis". (Accessed 16 September 2008). Available: http://www4.gsb.columbia.edu/null/download?&exclusive=filemgr.download&file_id=27378

Ehrenberg, G. and Smith, S. (1994) *Modern Labour Economics: Theory and Public Policy*, Fifth edition. New York: Harper Collins College Publishers.

El-Badry, M. (1961) "Failure of Enumerators to Make Entries of Zero: Errors in Recording Childless Cases in Population Censuses". *Journal of the American Statistical Association* 56(296):909-924.

Gangadharan, and Maitra, P. (2003) "Testing for Son Preference in South Africa". *Journal of African Economics* 12(3):371 – 416.

Garenne, M. (2004) "Sex ratios at birth in populations of Eastern and Southern Africa". *Southern African Journal of Demography* 9(1):91-96.

Gormick, J., Meyers, M. and Ross. K. (1996) "Public polices and employment of mothers: A cross-national study". *Luxembourg Income Study Working Paper Series,* Working paper no. 140.

Gronau, R. (1973) "The Effect of Children on Housewive's Value of Time". *The Journal of Political Economy* 81(2): S169 – S199.

Gronau, R. and Hamermesh, D. (2001) "The Demand for Variety: A Household Production Perspective." Working Paper 8509, NBER.

Gustafsson, S. and Worku, S. (2006) Marriage Markets and Single Motherhood in South Africa. *Tinbergen Institute Discussion Paper*, TI 2006-102/3. (Accessed May 2008). Available http://www.tinbergen.nl/discussionpapers/06102.pdf

Gustafsson, S. and Worku, S. (2008) "Teenage Motherhood and Long-Run Outcomes in South Africa". Presented at the African Econometric Society Conference, July 2008, Pretoria.

Iacovou, M. (2001) "Fertility and Female Labour Supply". *Working Papers of the Institute for Social and Economic Research*, paper 2001-19. Colchester: University of Essex.

James, D. (2001) "Land for the Landless: Conflicting Images of Rural and Urban in South Africa's Land Reform Programme". *Journal of Contemporary African Studies* 19(1):93-109.

Jensen, E. and Ahlburg, D. (2004) "Why Does Migration Decrease Fertility? Evidence from the Philippines". *Population Studies* 58(2):219 – 231.

Joshi, H. (2002) "Production, Reproduction, and Education: Women, Children, and Work in a British Perspective". *Population and Development Review* 28(3): 445-474.

Hallman, K., Quisumbing, A., Ruel, M. and de la Briere, B. (2002) "Childcare, Mother's Work, and Earnings: Findings from the Urban Slums of Guatemala City". Policy Research Division Working Paper No. 165.

Hsiao, C. (2003) "Chapter 1: Introduction". In: Hsiao, C. (2003) *Analysis of Panel Data*, second edition. Cambridge University Press. Econometric Society Monographs No. 34. pp.1-11.

Hunter, M. (2006) "Fathers without *amandla:* Zulu-speaking men and fatherhood". In: Richter, L. and Morrell, R. (eds) *Baba: Men and Fatherhood in South Africa.* Cape Town: HSRC Press pp. 99 – 109.

Kalule-Sabiti, I., Palamuleni, M., Makiwane, M. and Amoateng, A.Y (2007) Chapter 5: Family formation and dissolution patterns. In. Amoateng, A.Y. and Heaton, T.B. (2007) *Families and Housholds in Post-Apartheid South Africa: Socio-demographic perspectives.* Human Science Research Council Press: Cape Town. pp. 89-112.

Kaufman, B. and Hotchkiss, J. (2006) "Chapter 3: Labour Force Participation". In: Kaufman, B. and Hotchkiss, J. (2006) *The Economics of Labour Markets,* seventh edition. Thomson South Western. pp. 114-170.

Korenman, S. and Neumark, D. (1992) "Marriage, Motherhood, and Wages". *The Journal of Human Resources* 27(2):233-255.

Lam, D., Seekings, J. and Sparks, M. (2006) "The Cape Area Panel Study: Overview and Technical Documentation Waves 1-2-3". The University of Cape Town. (Accessed 7 August 2008). Available http://caps.psc.isr.umich.edu/documentation/overview

Leibowitz, A. and Klerman, J. (1995) "Explaining changes in married mothers' employment over time". *Demography* 32(3): 365 – 378.

Lloyd, C. (1991) "The Contribution of the World Fertility Surveys to an Understanding of The Relationship Between Women's Work and Fertility". *Studies in Family Planning* 22(3):144-161.

Magwaza, T. (2003) "Perceptions and experiences of motherhood: A study of Black and White mothers of Durban, South Africa". A *Journal of Culture and African Women Studies*: Issue 4.

Maitra, P. and Ray, R. (2003) "Resource Inflows and Household Composition: Evidence From South African Panel Data". *Journal of International Development* 15:1037-1047.

Makiwane, M. (1996) "Why is Fertility Declining in South Africa, Case Study of Mtombo, Eastern Cape". African Population Studies 11(1):200-213.

Marini, M. (1984) Women's Educational Attainment and the Timing of Entry into Parenthood, *American Sociological Review*, 49(4): 491-511.

Marteleto, L., Lam, D. and Rinchod, V. (2006) "Schooling and Early Childbearing in Urban South Africa". Research Report 06-610. Population Studies Centre: University of Michigan.

Mason, K. and Palan, V. (1981) "Female employment and fertility in Peninsular Malaysia: The maternal role incompatibility hypothesis reconsidered". *Demography* 18(4): 549 – 575.

May, J., Carter, M. Haddad, L. and Maluccio, J. (1999) "KwaZulu-Natal Income Dynamics Study (KIDS) 1993 – 1998: A longitudinal household data set for South African policy analysis". CSDS Working Paper No.21. School of Development Studies, University of KwaZulu-Natal.

McDaniel, A. and Zulu, E. (1996) "Mothers, fathers, and children: Regional patterns in child-parent residence in sub-Saharan Africa". *African Population Studies* 11:1–18.

Mencarini, L. (1999) "An analysis of Fertility and Infant Mortality in South Africa based on 1993 LSDS data". Paper presented for the Third African Population Conference "African Population in the 21st Century", Durban, South Africa, 6-10 December 1999.

Mincer, J. (1963) "Labour Force Participation of Married Women: A Study of Labour Supply". In: H. Gregg Lewis (ed) *Aspects of labour economics.* Princeton, NJ: Princeton University Press. Pp 63-105.

Mincer, J. and Polachek, S. (1974) "Family Investments in Human Capital: Earnings of Women". *The Journal of Political Economy* 82(2), Part 2: Marriage, Family Human Capital, and Fertility: S76-S108.

Mkhize, N. (2006) "Chapter 15: African traditions and the social, economic and moral traditions of fatherhood". In: Richter, L. and Morrell, R. (2006), *Baba: Men and Fatherhood in South Africa.* Cape Town: HSRC Press. pp 183 – 198.

Mlantsheni, C and Leibbrandt, M. (2001) "The Role of Education and Fertility in the Participation and Employment of African Women in South Africa". DPRU Working Papers No 01/54, September 2001.

Mostert, W.P and Hofmeyr, B.E. (1988) "Socioeconomic factors affecting fertility in the developing countries of the developing population groups in South Africa". *Human Sciences Research Council.*

Moultrie, T. and Dorrington, R. (2004) "Estimation of Fertility from the 2001 South Africa Census Data". CARe Monograph no.12, Centre for Acturial Research: University of Cape Town.

Moultrie, T. and Timaues, I. (2001) "Fertility and Living Arrangements in South Africa". *Journal of Southern African Studies* 27(2):207 – 223.

Moultrie, T. and Timaues, I.(2002) "Trends in South African Fertility between 1970 and 1998: An analysis of the 1996 Census and the 1998 Demographic Health Survey". A technical report of the Burden of Disease Research Unit and Medical Research Council.

Moultrie, T. and Timaues, I. (2003) "The South African Fertility Decline: Evidence from Two Censuses and a Demographic and Health Survey". *Population Studies* 57(3):265-283.

Mutedi, A. (2002) "Marital Status and the South African labour market. Statistics South Africa: Occasional Paper".

Mwabu, G. and Schultz, T. (2000) "Wage Premiums for Education and Location of South African Workers, by Gender and Race". *Economic Development and Cultural Change* 48:307-334.

National Income Dynamics Study (2008) "Welcome to the National Income Dynamics Study (NIDS)". (Accessed 10 November 2008) Available: http://www.nids.uct.ac.za

Noumbissi, A., Bawah, A. and Zuberi, T. (2005) "Chapter 6: Parental survival and residential patterns". In: Zuberi, T., Sibanda, A. and Udjo, E. (eds) (2005) *The Demography of South Africa,* ME Sharpe, Inc. pp. 160-177

Palamuleni, M., Kalule-Sabiti, I. and Makiwane, M. (2007) "Chapter 6: Fertility and childbearing in South Africa". In. Amoateng, A.Y. and Heaton, T.B. (2007) *Families and Housholds in Post-Apartheid South Africa: Socio-demographic perspectives.* Human Science Research Council Press: Cape Town. pp. 113 – 133.

Piras, C. and Ripani, L. (2005) "The effects of motherhood on wages and labour force participation: Evidence from Bolivia, Brazil, Ecuador and Peru". *Sustainable Development Department Technical Paper Series.* Washington, D.C.: Inter-American Development Bank.

Posel, D. (2001) "Who are the heads of household, what do they do, and is the concept of headship useful? An analysis of headship in South Africa". *Development Southern Africa* 18(5):651-670.

Posel, D. (2002) "A Review of Current Literature and Research on Migration in Southern Africa". (accessed 10 September 2008) available: http://www.queensu.ca/samp/migrationresources/Documents/Posel_review.pdf

Posel, D. and Casale, D. (2003) "What has been happening to internal labour migration in South Africa, 1993-1999?" *South African Journal of Economics* 71(3):455-479.

Posel, D. and Casale, D. (2006) "Chapter 10: Internal labour migration and household poverty in post-apartheid South Africa". In: Bhorat, R. and Kanbur, R. (2006) *Poverty and Policy in Post-Apartheid South Africa,* Cape Town: HSRC Press. pp. 351-365

Posel, D. and Devey, R. (2006) "Chapter 4: The demographics of fathers in South Africa: an analysis of survey data, 1993 – 2002". In: Richter, L. and Morrell, R. Baba: *Men and Fatherhood in South Africa*, Cape Town: HSRC Press. pp. 38 – 52.

Posel, D., Fairburn, J. and Lund, F. (2006) "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa". *Journal of Economic Modelling*, 23(4):836 – 853.

Posel, D. and van der Stoep, G. (2008) "Co-resident and Absent mothers: Motherhood and Labour Force Participation in South Africa". Paper presented at the AES conference, Pretoria, July 2008 and at the Conference on Income Distribution and the Family, Kiel (Germany).

Preston-Whyte, E. (2003) "Contexts of vulnerability: Sex, secrecy and HIV/AIDS" *African Journal of AIDS Research* 2(2):89-94.

Ramphele, M. and Richter, L. (2006) "Chapter 6: Migrancy, family dissolution and fatherhood". In: Richter, L. and Morrell, R. Baba: Men and Fatherhood in South Africa, Cape Town: HSRC Press. pp. 73 – 81.

Rosenzweig, M. and Wolpin, K. (1980) "Life-cycle labour supply and fertility: Causal inferences from household models". *The Journal of Political Economy* 88(2): 328 – 348.

Russell, M. (2004) "Understanding Black Households in South Africa: The African kinship and Western Nuclear Family Systems". Centre of Social Science Research, University of Cape Town. Working paper No. 97.

Schultz, P. (1990) "Testing the Neoclassical Model of Labour Supply and Fertility". *The Journal of Human Resources* 25(4):599-634.

Sibanda, A. and Lehloenya, G. (2005) "Race and Gender Gaps in Education". In: Zuberi, T., Sibanda, A. and Udjo, E. (eds) (2005) *The Demography of South Africa*. M.E. Sharpe Inc. pp. 253 – 265.

Simpson, J. (1995) "Pregnancy and the timing of intercourse". The *New England Journal of Medicince* 333 (23):1563-1565.

Statistics South Africa (2002). "Census 2001: Metadata. Introduction".
Pretoria: South Africa, Statistics South Africa. (Accessed 10 March 2008) Available: http://www.statssa.gov.za/census01/html/C2001metadata.asp .

Stycos, M. and Weller, R. (1967) "Female Working Roles and Fertility". *Demography* 4(1):210 -217.

Sudarkasa (2004) "Conceptions of Motherhood in Nuclear and Extended Families, with Special Reference to Comparative Studies Involving African Societies". *Journal of Culture and African Women Studies* Issue 5. (Accessed 6 March 2008). Available: http://www.jendajournal.com/issue5/sudarkasa.htm

Stock, J., Wright, J. and Yogo, M. (2002) "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments". *Journal of Business and Economic Statistics* 20(4):518-529.

Tanser, F., Hosegood, V., Barnighausen, T., Herbst, K., Nyirenda, M., Muhwava, W., Newell, C., Viljoen, J., Mutevedzi, T. and Newell, M. (2007) "Cohort Profile: Africa

Centre Demographic Information system (ACDIS) population-based HIV survey". *International Journal of Epidemiology* 37(5):956-962.

Tiefenthaler, J. (1997) "Fertility and family time allocation in the Philippines". *Population and Development Review* 23(2):377 – 397.

United Nations (2008) "World Population Prospects: The 2006 Revision – population sex ratios". (Accessed 9 September 2008) Available: http://data.un.org

University of Witwatersrand, The (2008) "About the Agincourt Health and Population Unit and HDSS". (Accessed 7 August 2008). Available: http://web.wits.ac.za/Academic/Health/PublicHealth/Agincourt/agincourt_hdss.htm

Townsend, N., Madhavan, S., Tolman, S., Garene, M. and Kahn, K. (2002) "Children's Residence Patterns and Educational Attainment in Rural South Africa, 1997". *Population Studies* 56(2):215 – 255.

Udjo, E. (1998) "Additional Evidence regarding Fertility and Mortality Trends in South Africa and Implications for Population Projections". (Accessed 1 December 2007). Available at http://www.statssa.gov.za/census01/census96/HTML/Metadata/Docs/fertility_udjo.htm

Van der Waal, C. (1996) "Rural children and residential instability in the northern province of South Africa". *Social Dynamics* 22(1):31- 54.

Vere, J. (2008) "Dragon Children: Identifying the Causal Effect of the First Child on Female Labour Supply with the Chinese Lunar Calendar". *The Oxford Bulletin of Economics and Statistics* 70(3):303-325.

Waldfogel, J. (1997) "The Effect of Children on Women's Wages". *American Sociological Review* 62(2):209-217.

Winter, C. (1999) "Women Workers in South Africa: Participation, Pay and Prejudice in the Formal Labor Market". World Bank Developmental Working Paper No. 19752. Washington, D.C.

Willis, R. (1973) "A New Approach to the Economic Theory of Fertility Behaviour". *The Journal of Political Economy* 81(2):S14-S64.

Wooldridge, J. (2003) "Chapter 15: Instrumental Variables Estimation and Two-stage Least Squares". In: *Introductory Econometrics, A Modern Approach, second edition.* South-Western Publishing.

Wong, R. and Levine, R. (1992) "The effect of household structure on women's economic activity and fertility: Evidence from recent mothers in urban Mexico". *Economic Development and Cultural Change* 41(1):89 – 102.

Zulu, E. and Sibanda, A. (2005) "Chapter 8: Racial Differences in Household Structure". In: Zuberi, T., Sibanda, A. and Udjo, E. (eds) (2005) *The Demography of South Africa* M.E. Sharpe Inc: Cape Town. pp. 218 – 252.

Zuberi, T. and Sibanda, A. (2005) "Chapter 10: Migration and Employment". In: Zuberi, T., Sibanda, A. and Udjo, E. (eds) (2005) *The Demography of South Africa* M.E. Sharpe Inc: Cape Town. pp. 266 – 296.

Data Sources:

South African Labour and Development Research Unit, *Project for Statistics on Living Standards and Development*, 1993. University of Cape Town.

Central Statistics Services, *October Household Survey*, 1995. Pretoria

Central Statistics Services, *October Household Survey*, 1997. Pretoria

Statistics South Africa, *General Household Survey*, 2002. Pretoria

Statistics South Africa, *1996 Population Census*. Pretoria

Statistics South Africa, *2001 Population Census*. Pretoria

University of KwaZulu-Natal, *The KwaZulu-Natal Income Dynamics Study*.

UNIVERSITY OF
KWAZULU-NATAL

RESEARCH OFFICE (GOVAN MBEKI CENTRE)
WESTVILLE CAMPUS
TELEPHONE NO.: 031 – 2603587
EMAIL: ximbap@ukzn.ac.za

19 JUNE 2008

MS. G VAN DER STOEP (203503108)
ECONOMICS

Dear Ms. van der Stoep

ETHICAL CLEARANCE APPROVAL NUMBER: HSS/0176/08M

I wish to confirm that ethical clearance has been approved for the following project:

"Identifying motherhood and its economics costs: An analysis of survey data in South Africa"

PLEASE NOTE: Research data should be securely stored in the school/department for a period of 5 years

Yours faithfully

Pximba
.....
MS. PHUMELELE XIMBA

cc. Supervisor (Prof. D Posel)
cc. Ms. J Mazibuko

UNIVERSITY OF
KWAZULU-NATAL

13 December 2008

Supervisor's report

Candidate: Gabrielle van der Stoep

Research Master's: Identifying motherhood and its effects on female labour force participation in South Africa: An analysis of survey data

Supervisor: Prof Dori Posel

I have read the near final draft of the thesis written by Ms Gabrielle van der Stoep and declare that it is presented in a form ready for examination.

I started supervising Ms van der Stoep's thesis in 2007. Over the past two years, we have met regularly to discuss the study, and to work through conceptual, data and estimation issues.

Ms van der Stoep has worked consistently, and with great dedication on her thesis. I have now read at least two drafts of each chapter of the thesis where I provided extensive comments for revision and clarification. Ms van der Stoep was always responsive to these suggestions and recommendations, and she frequently took an extra step in extending the analysis.

Without making any suggestion as to the mark it should receive, I believe that this thesis is a good reflection of the student's ability.

Yours sincerely

Prof Dori Posel
School of Development Studies
UZKN
posel@ukzn.ac.za

UNIVERSITY OF
KWAZULU-NATAL

13 December 2008

I declare that the thesis written by Ms Gabrielle van der Stoep is written at an acceptable standard of English.

Yours sincerely,

Prof Dori Posel
School of Development Studies
UZKN
posel@ukzn.ac.za