The causal effect of family size on mothers' labor supply: evidence from Reunion Island and mainland France

Idriss Fontaine*

September 23, 2016

FIRST DRAFT
DO NOT QUOTE

Abstract

I provide new evidence about the negative impact of family size on mothers' participation to the labor market for the particular region of Reunion Island and mainland France. Based on an instrumental variables (IV) model of labor supply, exploiting two exogenous increases of fertility through sex mix composition and multiple births, I find that the negative causal impacts are quite different. For Reunionese mothers the presence of a third child decreases the probability to participate of about 15 percentage points, while, the same effect is two times less important for French mothers. This finding reinforces the idea that family policies should use specific tools to facilitate the return to the labor force of Reunionese mothers.

Keywords: Labor supply, family size, instrumental variables, Reunion Island
JEL classifications: J13, J22

*Université de La Réunion - 15 Avenue René Cassin, Saint-Denis 97400, La Réunion - Centre d'Economie et de Management de l'Océan Indien (CEMOI) - email: idriss.fontaine@univ-reunion.fr / idriss-fontaine@hotmail.fr - Tel: 0692 83 82 03.

1 Introduction

In the Reunion Island[1] the participation rate to the labor market of women has experienced an important increase. In four decades, female participation increased by 30 percentage points, starting from 30% in 1974 to reach 60% in 2012. During the same period the average number of children per women decreased dramatically since it declines from 6 children per women to 2,4. These two major trends about female participation and fertility behaviors are striking features also shared by French women since the post-World War II[2]. Nonetheless, a snapshot of women activity in 2012 shows that the labor supply of Reunionese women continues to be lower of approximately 10 percentage points comparatively to its French mainland level. Furthermore, the fertility indicators still indicate that a Reunionese woman has more children than in France. The main objective of this paper is to study in what extent the discrepancy in term of women fertility can be an explanation of the different level of female activity among the specific region of Reunion and mainland France.

The concrete measure of the relationship between family size and women labor supply is not straightforward. Although, *prima facie* the correlation between them appears to be negative, it is quite hard to disentangle a causal relationship for at least three reasons. First, the two behaviors are influenced by common factors. Second, both fertility and participation decisions are jointly determined and it is difficult to determine whether the former is a cause or an aftermath of the latter. Third, adverse selection complicates the identification of a causal impact. As a consequence of these three shortcomings, simple models as Ordinary Least Square (OLS, thereafter) suffer from endogeneity bias. In order to identify a causal effect of the number of children on mother's participation an exogenous source of variation in family size is needed. In this respect, I follow Angrist and Evans (1998) by estimating an instrumental variable model of labor supply. The endogenous variable of fertility is instrumented by the sex mix of the two eldest and by an indicator of twin birth at second motherhood. Similar strategy of instrumentation has been used in many works as Cruces and Galiani (2007) for Argentina and Mexico, Frenette (2011) for Canada or Caceres-Delpiano (2012) for a sample of developing countries.

The identification of a causal relationship should not be only of interest of academic researchers. In this respect, this paper may be of interest of policy designers. It is crucial to understand the reason explaining the decrease in female participation with the number of children. For instance, if this decline is not the aftermath of female fertility but rather a consequence of other factors as the education level, policies aiming at reconciling professional

[1]For simplicity, in the rest of the paper I use the short term of Reunion to designate the Reunion Island.

[2]The decline in fertility is less pronounced in France since its level amounted to 2,5 children per women in 1970 and reached a value of 2 in 2012.

and private life will be not efficient[3]. In contrast, if fertility has a causal impact on women activity, then such policies will reach its target and may be efficient. In the special case of Reunion, the interest should be even higher. If the family size has a more negative impact on women participation in Reunion than in France, policy makers should take into account this specificity to adapt family policies to this particular context.

This article makes several contributions to the literature. Firstly, it is the first to use the New Census of the French Population to estimate this causal relationship. This database is interesting in several respects. It contains a lot of variables, covers the entire French territory and it is the sole which allows me to retrieve a large sample in the case of Reunion. The latter element is non-trivial because other datasets do not allow having a sufficient number of observation for Reunion. Secondly, this article provides an update of evidence for the case of France since the latter paper dealing with this issue uses relatively old data spanning the 1990-2002 period. Thus, by using data spanning the 2008-2012 period, I can investigate if the causal impact is varying in France. Put differently, is the causal impact of family size on female participation the same between 1990-2002 and 2008-2012 periods? Thirdly, my paper is the first to provide evidence for Reunion. This region is specific for several reasons. It has the same institutions than in France but its economy is characterized by insularity and its labor market by high level of unemployment rate.

Simple OLS estimators confirm a negative correlation between labor supply outcome and motherhood. Interestingly, the correlation is of same magnitude for the two regions studied. IV models using sibling sex composition do not provide significant results for the Reunionese case. To circumvent this pitfall, I favor a combined use of the twin births and sex composition of the two eldest as instruments. This strategy increases the population of treated and the precision of estimations. My final conclusion is as follow: the magnitude of the causal effect appears to be quite different across the two samples studied. Thus, the presence of a third child implies a reduction in mothers' labor supply probability of about 15 percentage points in Reunion. For French mothers, the causal impact of fertility on participation is two times less important and contributes to lower the probability of female participation of only 7,5 points. In both regions the causal effect is significantly negative but it remains substantially high for Reunionese mothers. This finding reinforces the idea that women childbearing is one of the possible candidates to explain the lower female participation rate in Reunion. My findings indicate that policy makers may increase female labor supply by implementing policies facilitating the reconciling between professional career and fertility. However, my evidence suggests that the negative impact of family size on female activity changes over time. With a similar strategy than me, Moschion (2009) finds that the negative influence of

[3]Child care policy may be an example of such policies.

fertility on participation amount to 19 points in mainland France. In contrast to her, and based on census data spanning the 2008-2012 period, I show that the negative impact of the number of children on French female participation declined.

The rest of this paper is structured as follow. Section 2 serves as preliminary analysis. It presents the literature and the Reunionese context. Section 3 describes the dataset and presents the econometric framework. In the section 4 the main results of the paper are presented. Section 5 discusses the results. Finally, the next section provides some concluding remarks.

2 Preliminary

2.1 Literature background

Simple statistic descriptives still show that on average women with children experience worse labor market outcomes than childless women[4]. Their chances to participate are lower and once in employment their chances to take a part time job are higher and their earnings lower (Korenman and Neumark (1992), Waldfogel (1998), Blundell et al. (2013)). These descriptive elements suggest that the correlation between women fertility and their labor supply is probably negative. Nonetheless, correlation should not be systematically understood as a causal relationship. In the case of the impact of fertility on female participation three arguments complicate the interpretation of direct evidence. Firstly, the two phenomenon may be explained by common factors. For instance, the education level of mothers may influence their career opportunities but also their childbearing behavior. Thus, it is necessary to add control variables in the model in order to mitigate this concern. Secondly, the fertility and the labor supply decisions are jointly determined and it is imperative to disentangle in what direction the causal effect operates. This is called the reverse causality problem. In particular, one can claims that depending on her career opportunities a woman may choose to diminish her desired number of children. Conversely, another one can argues that a woman with a large family may choose to reduce her activity or to withdraw from the labor force. Thus, it is quite difficult to say whether fertility is the cause or the consequence of the participation behavior. Thirdly, the adverse selection problem also impedes the identification of a causal effect. Adverse selection refers to the situation in which women choosing to have a large family have weaker labor force attachment than women choosing to remain childless. This is potentially problematic because it implies a bias in the correlation: women with children have a worse situation on the labor market whatever the presence of children. As a result

[4]See table 2 for an illustration.

the assessment of a causal impact is complicated by the presumed endogeneity of fertility in the estimation of the labor supply equation.

To distinguish causal evidence from adverse selection, empirical studies propose to estimate the equation of interest by the instrumental variable method. The purpose is to find variables highly correlated with the endogenous fertility variable but not directly related to the labor market outcome. The instruments should create an "exogenous" fertility shock, randomly assigned in the sample such that the affectation is similar to a natural experiment. Different sources of exogenous variations of fertility were used, the first being the incidence of multiple births (Rosenzweig and Wolpin, 1980). The twin birth (at first, second, third or fourth motherhood) is probably a good instrument in the extent to which it cannot be anticipated, it is almost not related to maternal characteristics[5] and it provokes an exogenous shift in fertility of the desired number of children[6]. Caceres-Delpiano (2012), with a sample of mothers from developing countries, shows that women with twins at first birth are in a worse situation that women with singletons, the probability of employment being 3 points lower. More recently, Li et al. (2015) find that fertility induced by multiple births does not affect significantly female participation to the labor market in rural China. Since the pioneer contribution of Angrist and Evans (1998), another trend of the literature consists in the exploitation of parental preferences for a mixed sibling sex composition as a natural experiment[7]. Since parents with two eldest of same gender have a higher probability to have a third additional child, and assuming that sex composition is randomly assigned, a variable indicating the sex composition of the first two children is probably a good instrument for fertility. This identification strategy has been followed by several authors and for quite different countries[8]. On the overall, it appears that an exogenous shock of fertility induced by the event of non-mix gender deteriorates female labor market outcomes. However, the magnitude of the causal effect varies: from approximately less than -10 points in the U.S., Canada or Argentina to -20 points in France. Furthermore, the causal impact is not significant in Sweden, Great Britain and Chile.

[5]This statement must be mitigated by at least two stylized facts. Firstly, it is known that older women are more likely to have twin. Secondly, multiple births are more frequent for women who undergo fertility treatment.

[6]There is a prosperous and too long to be exhaustively quoted literature which use the event of twin birth as a natural experiment.

[7]Recently, other instrumental variables have been proposed in the literature. Agüero and Marks (2008) use infertility shock whereas Lundborg et al. (2014) propose the use of in-vitro fertilization success as natural experiments. These two identification strategies provides consistent estimates of the impact of fertility on female labor supply. However, to be replicated these identification strategies need specific database which contain private and sensitive information.

[8]A non-exhaustive list: Cruces and Galiani (2007) for Argentina and Mexico, Frenette (2011) for Canada, Hirvonen (2010) for Sweden, Moschion (2009) for France.

2.2 The reunionese context

Reunion was a former French colony. After the World War II the island obtained the status of French department, while in 1982 it became a French oversea region. At the middle of the sixties Reunion began its economic take off. During 25 years between 1974 and 1999 the economic growth amounted to an average of 5% per year. This strong economic growth was accompanied by an increase in the Human Development Index (HDI) which is approximately equal to 0,87 in 2004 (Goujon, 2008). However, from economist and politician point of view the most striking feature characterizing Reunion is not its relatively high level of economic growth but rather its specific labor market with a high level of unemployment. Thus, in 2012 the unemployment rate was approximately equal to 28%. Although, this specific feature should attract a great deal of attention, it should not hide other important changes of the labor market of Reunion. For instance, the economic and development expansions have changed female behavior in several aspects. The female participation rate doubled in 40 years indicating that Reunionese women are more likely to supply their labor. Furthermore, women fertility declined dramatically. Despite these concomitant evolutions, the participation rate of women is lower in Reunion and the average number of children higher than in France. Furthermore, the Reunion case is interesting because it share the same institutions than in France. Family policies are very close across the two regions. In this particular context, the identification of a causal relationship between fertility and female participation may reveal that policies should be adapted to take into account the specificity of the oversea region.

3 Empirical methodology

3.1 Data and descriptive statistics

The information about female labor supply and their fertility are extracted from the New Census of the French Population [9] (NCFP, thereafter). Since 2004 the French Census is a survey and a global representativeness of the French population is achieved by stacking five consecutive annual surveys. I use the 2010 version of the NCFP which is therefore composed by the 5 waves starting from 2008 to 2012. For France the raw data contain 19,716,384 observations while for Reunion there are 318,142 observations. The use of this dataset is crucial because it is the sole allowing to have a very large sample for Reunion, a necessary condition to have robust estimations. The NCFP is also interesting because it contains a lot

[9] In French: *"Le Recensement Rénové de la Population"*

of socio-demographic variables such as the age, the education and the labor market state[10]. Most prominently, it gives a set of information about the family structure (sex of children, date of birth etc.) which is of first interest for the topic of this article. Differently from Angrist and Evans (1998) and for the sake of robustness of the empirical results[11], I restrict the sample to mothers aged between 21 and 40 years old with at least two children. As a result, I analyze the effect of a third birth on mothers' labor supply. After these restrictions for France I have a sample of 878,215 mothers and for Reunion a sample of 19,957 mothers.

Table 2 documents a summary of descriptive statistics about the sample of mothers with at least 2 children for both France and Reunion. In the restricted sample, French women are older than their counterparts. In contrast, Reunionese mothers have their first child earlier, 22 years old in Reunion against 25 years old in France[12]. The levels of female qualifications are very different across the two samples. For Reunion women are more often unskilled and less often skilled than in France. This finding is consistent with official figures. In terms of labor supply, the labor market participation rate is around 80% in France. The female participation is lower in Reunion for which it is approximately equal to 70%. In the bottom part of table 2, I provide a set of figures about fertility. The dummy variable indicating if the mothers have more than 2 children is of primary interest in the labor supply equation I aim to estimate, since it corresponds to the presumed endogenous variable. Thus, around 40% of mothers with at least two children have at least one additional child for women living in Reunion. The corresponding figure for French mothers is 32,7%. As a result, the average number of children is lower in France. With regard to the instrumental variables (first two children of same sex and its alternatives, twins at second birth), the figures of the table do not reveal huge discrepancies between the two samples. In both France and Reunion, 50% of the first two children have the same sex and slightly above 1% of second birth correspond to multiple births.

3.2 Econometrics

As argued by Angrist and Evans (1998) the OLS estimates do not provided unbiased estimators of the labor supply equation. To circumvent the potential pitfall of endogeneity in the estimation of the participation equation, instrumental variables (IV) are used in the

[10]Here, it is important to note that the definitions of labor market states (employment, unemployment and not in the labor force) are different from the ILO's standards. Actually, each individual reports "spontaneously" his labor position.

[11]I have checked the robustness of the result to this restriction.

[12]These figures about the average age at first birth is quite different of the official statistics of the French National Institute os Statistic and Economics Studies (FNISES). The official average age at first birth amount to 28 years old in France and to 25,6 years old in Reunion. The sample restrictions imposed entirely explain this discrepancy

	France	Reunion
Socio-demographic		
Age	35,46	34,54
Age at first birth	25,22	22,43
No degree	0,153	0,400
< Bac	0,274	0,271
Bac	0,206	0,163
Bac + 2	0,199	0,088
> Bac +2	0,168	0,078
Participation rate	0,806	0,699
Participation rate with 2 children	0,867	0,767
Participation rate with at least 3 children	0,681	0,595
Fertility		
Number of children	2,43	2,62
More than 2 children	0,327	0,394
2 boys first	0,265	0,268
2 girls first	0,240	0,238
First two children of same sex	0,505	0,506
Boy first	0,514	0,520
Twins at the 2nd birth	0,016	0,014

Table 1: Summary statistics of mothers with at least two children
Source: NCFP, author's own calculations

empirical model. More specifically, along the lines of Angrist and Evans (1998), Agüero and Marks (2008) or Caceres-Delpiano (2012) (among other), I estimate a Two-Stage Least Squares (2SLS) linear probability model. The 2SLS estimator has several practice advantages for the problem in hand. In contrast to maximum likelihood model, it does not specifies the distribution of error terms. It is relatively simple to implement and, as argued by Angrist and Pischke (2009), when it comes to approximate marginal effects it works well. Then, this model does not specify any type of distribution for the endogenous variables. They may be censored, discrete or continuous (Lewbel et al., 2012). The sole assumption required for the estimation is that the IV are uncorrelated with the error term of the regression. Put differently, instruments must be themselves exogenous. Unfortunately, there is no direct test of the exogeneity hypothesis. In order to mitigate this concern, remind that the exogeneity hypothesis is common to all IV model. Finally, observe that all the instruments used in this paper are dummy variables. Implicitly, it is assumed that the sample is split into two sub-groups. The first sub-group of mothers has an incentive to increase its fertility while the

second one has not.

The first stage consists in estimating the fertility equation linking the variable "more than two children" x_i, to the instruments z_i and other socio-demographic covariates w_i:

$$x_i = \pi_0^{'} w_i + \gamma z_i + \eta_i \tag{1}$$

with η_i the residuals of the first stage estimation. Since I will estimate several models, the variable indicating the presence of a third childbearing will be instrumented, either by a dummy equal to 1 if the sex of the second child is the same than the first child (0 otherwise), either by a dummy indicating if the mothers have had a multiple births for her second motherhood, or by a combined use of these instruments. Note that the "same-sex" instrument can be divided into two others variables: two eldest boys and two eldest girls. As it is usual in this literature, and in order to improve the precision of the estimations, I add a list of exogenous variables to the regressions. Thus, the vector w_i contains a set of socio-demographic variables, such as, the age, the age at the first birth, the gap between the first two childbearing, the education level (in 5 levels), 5 annual fixed effects, a dummy indicating the sex of the first child and a dummy indicating if the mother was born abroad.

After the estimation of the first stage equation, the labor supply equation can be estimated:

$$y_i = \alpha^{'} w_i + \beta x_i + \varepsilon_i \tag{2}$$

where y_i is the labor participation variable equal to 1 if the woman supply her labor and 0 if she is out of the labor force, $\alpha^{'}$ the vector of coefficient for the control variables. The coefficient β is of main interest. It tells us in what extend the probability of participation varies for a marginal variation (from 2 to 3) of the number of children only for mothers affected by the instruments (those with two eldest of same gender or those who experience a multiple birth at their second motherhood). As argued by Angrist and Imbens (1995), β should be interpreted only as a local average treatment effects (LATE)[13].

4 Results

4.1 First stage results

Estimations of the first stage fertility equation which have the indicator "more than two children" as dependent variable are presented in table 2. These results should be analyzed

[13] β corresponds to the average treatment effect (ATE) only if the population affected by the instrumental variables is representative of the overall population.

finely since it provides some clues about the appropriateness of the instruments. First, all estimated coefficients reported in the table are positive and statistically significant at 1% level. This indicates a non-negligible correlation between the instrumental variables and the endogenous fertility variables. Stock and Yogo (2005) propose several criteria for selecting minimum threshold for the F statistic of first stage equations. Concretely, an F statistic above 10 suggests a non-weak instrument. Here, all F statistics are largely above 10, especially for the French case. This is a direct consequence of the large sample size. The correlation between the instruments and the endogenous variables combined with the above criteria suggest that my second stage estimations will not be affected by the concern of weak instrument. Second, observe that the values of the estimated coefficient are modified only slightly when the set of covariates are included in the estimation. This finding is important because it indicates no strong correlation between them. Thus, the influence of the instruments (on the endogenous fertility variable) cannot be attributed to any noise implying these observed variables. These two characteristics gives some arguments justifying the use of the "same sex" and "twin" variables as instruments.

Now, let me turn to a detailed analysis of the first stage estimation. On the overall, the regressions confirm the idea that mothers prefer to have children with different gender. In a model without any covariates, the probability to have a third child when the first two eldest are of same sex is higher of 5,7 percentage points in Reunion against 4,3 points in France. When the "same-sex" variable is split into two separate variables ("two boys" and "two girls") the finding is similar. For both France and Reunion the probability to have an additional child is slightly higher when the first two eldest are daughters. However, the discrepancy between the two estimated coefficients is too weak to conclude on a sex preference of children in favor of boys. In the two last columns a set of socio-demographic is added to the regressions and allows me to have more precise estimates of the effect of gender sibling sex on further childbearing. As a result, the estimation suggests that in Reunion mothers who have had two eldest of same sex are 5,3 points more likely to have a third motherhood than mothers with two eldest of different sex. The corresponding figure amounts to 4 points in France. Concerning the twin birth instrument, estimations including covariates suggest that mothers having twin at their second motherhood are 68 percentage points more likely to have at least a third child than other mothers with at least two children in France. The same effect is slightly lower in Reunion since it is approximately equal to 63 points. This finding is not surprising. It is a direct consequence of the higher fertility in Reunion. On the overall, the different estimations confirm that mothers in Reunion are much more likely to have further childbearing which correspond to an expected finding.

	(1)		(2)		(3)		(4)		(5)		(6)	
	Fr	**Run**	**Fr**	**Run**	**Fr**	**Run**	**Fr**	**Run**	**Fr**	**Run**	**Fr**	**Run**
Same-sex	0,043 *** (0,001)	0,057 *** (0,007)	–	–	–	–	0,040 *** (0,001)	0,053 *** (0,005)	–	–	–	–
Two eldest boys	–	–	0,042 *** (0,001)	0,055 *** (0,001)	–	–	–	–	0,040 *** (0,001)	0,047 *** (0,008)	–	–
Two eldest girls	–	–	0,044 *** (0,001)	0,059 *** (0,009)	–	–	–	–	0,040 *** (0,001)	0,059 *** (0,008)	–	–
Twins - 2	–	–	–	–	0,680 *** (0,004)	0,615 *** (0,029)	–	–	–	–	0,685 *** (0,003)	0,631 *** (0,024)
F statistic	1833	67,86	917	34,05	30930	459,7	19450	608,0	18160	567,6	23130	670,2
Covariates	**No**						**Yes**					

Table 2: The effects of the gender composition on a third childbearing: first stage estimation.

Source: NCFP, author's own calculations

Notes: Fr refers to France while Run refers to Reunion Island. Standard errors are reported in parenthesis. Significant levels: * 10%; ** 5%; *** 1%. Sample size: for France 871,215; for Reunion 19,957. The estimations are based on a sample of mothers aged between 21 and 40 years old which have at least two children. Column (1): regression of the fertility variable on the sole variable indicating if the first two children are of same sex. Column (2): regression of the fertility variable on two variables one indicating if the first two children are boys and another one if the first two children are girls. Column (3): regression of the fertility variable on the sole variable indicating the event of multiple births at second motherhood. Column (4), (5) and (6): respectively repeat the models of column (1), (2) and (3) but with a set of control variables. The control variables are: age, age at first birth, the gap between the first two motherhood, a dummy indicating the sex of the first child, the education level (in 5 levels), 5 annual fix effects and a dummy indicating if the mother was born abroad.

4.2 Second stage results

The second stage estimation linking female participation with fertility and control variables are displayed in table 3. Due to space limitations, only the coefficients of the presumed endogenous variable "more than two children" are reported for separate regressions in which the instruments used is varying. In addition, since the results are insensitive to the inclusion of control variables (not shown here) I report only the estimates including controls. As shown in column 2 of table 3 the OLS estimates suggest a negative correlation between the third additional child and female participation in both samples. According to these estimates the presence of a third child reduces the probability of participation of about 19 percentage points. The magnitude of the negative effect is very close for the two regions. However, as indicated in the last section, simple OLS models suffer from endogeneity bias.

The next two columns (2SLS (1) and 2SLS (2)) present the first IV estimates in which the fertility indicator is instrumented by the gender composition of the two eldest. For France, the impact of the number of children on women fertility appears to be non-negligible. Thus, the presence of a third child seems to have a negative impact on labor participation by reducing it by 15 points. For Reunion the coefficient is lower but indistinguishable from 0. In particular, the estimated coefficient is equal to -9,3 but it is fairly imprecise with a standard error of 11,7 percentage points. The finding is similar when the instrument "same-sex" is divided according to gender. The phenomenon which operates here for the Reunion case is not uncommon in the literature and one characteristic seems to explain it: the sample size. All else being equal, the smaller the sample size is, the greater the standard error of the estimated coefficient is. This is especially true when IV estimators with one or several endogenous regressors are implemented. For example, with a sample of about 150,000 mothers Angrist and Evans (1998) find a significant impact of fertility on labor market outcomes with a standard error equal to 2 percentage points. Moschion (2009) also finds, with a smaller sample of 70,000 French mothers, a significant but fairly imprecise impact (with a standard error of 10 points) of family size on labor outcomes when the sex composition of the first two children is used as instrument. Finally, based on data from Chile, Lopez de Lerida (2005) finds a non-significant impact. However, she worked with a relatively small sample including only 7,000 mothers. By containing approximately 20,000 mothers against 897,000 for France, the Reunion case is consistent with this empirical finding.

The dummy variable "twins at the second birth" is also a valuable instrument for the endogenous fertility variable. In column 4 (2SLS (3)) I provide the corresponding results. The use of multiple births as instrument changes the precision and the magnitude of the negative effect. A possible explanation for the discrepancy in estimates between the set of instruments if that they consist in different treatment: a desired change versus a non-desired

in family size. In other words, the effect of fertility is not the same across mothers who have had a non-mix gender for her first two childbearing and those who have experienced a multiple second birth. Thus, the treatment effect is heterogeneous. Furthermore, observe that for France the negative impact is lower when the endogenous fertility variable is instrumented by the events of multiple births. In order to explain this discrepancy, Angrist and Evans (1998) argue that when the second birth corresponds to twins the third child is necessarily older than other third child. As a consequence, in the latter case the effect of fertility on labor supply is larger because the child is younger forcing mothers to allocate more time for children education. A possible alternative to limit this pitfall consists in a joint use of the two instruments. As argued by Angrist (2004) and Ebenstein (2009), this strategy increases the size of the complier population leading ultimately to results closer to the ATE. The corresponding regression can be found in the last column of table 3. Since, the precision is higher when the two instruments are used together this empirical framework is the one I favor to provide my final conclusions. Thus, with this model the negative effect of the third additional child is significantly negative for both samples and the precision of the estimates are correct. However, the magnitude of the effect is different. A Reunionese mother is much more likely to not supply her labor when she has at least 3 children than a French mother. Most prominently, with a third additional child, the probability of participating decreases by 15 percentage points for reunionese mothers. In France, the same negative impact of fertility on female labor supply is halved and amounts to 7,5 points.

	OLS	**2SLS (1)**	**2SLS (2)**	**2SLS (3)**	**2SLS (4)**
Instrument(s)	–	*Same-sex*	*Same-sex divided*	*Twins-2*	*Same-sex and Twins-2*
France	-0,192 *** (0,001)	-0,150 *** (0,002)	-0,148 *** (0,002)	-0,071 *** (0,005)	-0,075 *** (0,001)
Reunion	-0,185 *** (0,007)	-0,093 (0,117)	-0,108 (0,117)	-0,154 *** (0,041)	-0,147 *** (0,039)

Table 3: The impact of fertility on labor market participation of mothers: two stage estimation

Source: NCFP, author's own calculations

Notes: Standard errors in parenthesis. Significant levels: * 10%; ** 5%; *** 1%. Sample size: for France 871,215; for Reunion 19,957. The estimations are based on a sample of mothers aged between 21 and 40 years old which have at least two children. In each regression only the coefficient estimated of the variable "more than two children" is documented. Column (1) "OLS": regression of the participation variable on the variable "more than two children". Column (2) "2SLS (1)": regression of the participation variable on the variable "more than two children" instrumented by a dummy indicating the mix of the two eldest. Column (3) "2SLS (2)": regression of the participation variable on the variable "more than two children" instrumented by two dummies (one for boys and one for girls) indicating the mix of the two eldest. Column (4) "2SLS (3)": regression of the participation variable on the variable "more than two children" instrumented by a dummy indicating the presence of twins at second birth. Column (5) "2SLS (4)": regression of the participation variable on the variable "more than two children" instrumented by two dummies one indicating the presence of twins at second birth and another one indicating the mix of the two eldest. All regressions include a set of control variables. The control variables are: age, age at first birth, the gap between the first two motherhood, a dummy indicating the sex of the first child, the education level (in 5 levels), 5 annual fix effects and a dummy indicating if the mother was born abroad.

5 Interpreting the evidence

5.1 The French evidence

For France, my final empirical model can be seen as an update of previous evidence. Thus, Moschion (2009) finds that an additional third childbearing diminishes the labor supply probability of about 19 percentage points. My results indicate a lower negative effect, but one characteristic is different across the two studies. Moschion (2009) constructs her sample from the French Labor Force Survey and the last one provides the ILO definitions of labor market states. In contrast, starting from the NCFP I can only use a declarative notion of the three states[14]. However, observe that the samples used across the two works do not cover the same period: 1990-2002 for Moschion (2009) against 2008-2012 in this paper. As a consequence, a temporal effect should not be rejected. In one of her appendices, Moschion (2009) also estimates the causal impact of fertility on female participation with the gender mix as an instrument from the 1990 French Census. In this context, she finds that a third motherhood lowers the probability of participation of about 27 percentage points. With data and estimation method directly comparable, (column 2SLS(1) of table 3) I find a negative impact of 15 points. This evidence combined with the main finding of this paper reinforce the idea that the impact of family size on female participation diminishes over time. Thus, it is entirely possible that the trade-off between fertility and labor supply changes over time inducing, with the evidence provided here, that French mothers may have a higher preference for their professional career.

5.2 The reunionese evidence

For the specific case of Reunion evidence is new and suggests that the negative effect of family size on female participation is still high. Two possible explanations can be invoked to explain this fact. First, the causal relationship estimated indicates that when a Reunionese mother has several children it is more difficult for her to rejoin the labor force and consequently to take a job. Second, it is also possible that the higher negative impact in the Reunionese case is a direct consequence of the labor market environment. In a context of high structural unemployment level the labor market is tightened from the worker side (comparatively to open vacancies there is an important number of unemployed). As an unemployment spell has a lower probability to be successful, the opportunity cost of having an additional third child is probably lower in Reunion. Thus, a Reunionese mother is more likely to withdraw

[14]The discrepancies between the two concepts may be not negligible for the inactivity state. Specifically, with the declarative concept, a normal individual may have some difficulties to distinguish the unemployment spells and the inactivity spells.

from the labor force to devote her time to the education of children.

The exercise conducted in this paper reveals that discrepancy in term of causal impact may be depth even if the institutions across the two regions are close. Of course, different cultural norms about the role of mothers in children's education could contribute to this discrepancy. Regardless of the reason explaining why the causal impact is higher in Reunion than in mainland France, this fact should attract a great deal of attention of policy makers. In terms of policies, my empirical evidence indicates that specific tools should be establish in Reunion to encourage Reunionese mothers to supply their work and to facilitate the reconciling between motherhood and professional career.

6 Concluding remarks

This paper is the first to provide empirical evidence on the causal relationship of fertility on mothers' participation for the specific French oversea region of Reunion. It also gives an update of this effect for mainland France. Based on the pioneer strategy of Angrist and Evans (1998) using an instrumental variables estimator which exploits two random sources of fertility, I find that the magnitudes of the effects are quite different across the two regions. In particular, having at least three children pushes down the probability to participate of about 15 percentage points for Reunionese mothers. The same estimate for France is lower and amounts for 7,5 points. On the overall my estimates suggest that the negative causal effect is reduced in France comparatively to the 1990 decade. For the specific region of Reunion, the empirical evidence indicates that the presence of several children is still a major obstacle to reconciling professional and family responsibilities.

This study is a first step for the understanding of the interaction between childbearing and participation in Reunion. Nonetheless, depending on the public authorities, I share the idea that further researches should focus on the evaluation of family policies, notably through the lens of the return to labor force of Reunionese mothers.

References

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

Agüero, J. M. and M. S. Marks (2011). "Motherhood and Female Labor Supply in the Developing World Evidence from Infertility Shocks". *Journal of Human Resources* 46(4), 800–826.

Angrist, J. and S. Pischke (2009). Mostly Harmless Econometrics: An empiricist's Companion. Princeton: Princeton University Press.

Angrist, J. D. (2001). "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice". *Journal of Business and Economic Statistics* 19(1), 2–16.

Angrist, J. D. (2004). "Treatment Effect Heterogeneity in Theory and Practice". *The Economic Journal* 114(494), C52–C83.

Angrist, J. D. and W. N. Evans (1998). "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size". *The American Economic Review* 88(3), 450–477.

Angrist, J. D. and G. W. Imbens (1995). "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity". *Journal of the American Statistical Association* 90(430), 431–442.

Blundell, R., A. Bozio, and G. Laroque (2013). "Extensive and Intensive Margins of Labour Supply: Work and Working Hours in the US, the UK and France". *Fiscal Studies* 34(1), 1–29.

Caceres-Delpiano, J. (2012). "Can We Still Learn Something From the Relationship Between Fertility and Mother's Employment? Evidence From Developing Countries". *Demography* 49(1), 151–17.

Cruces, G. and S. Galiani (2007). "Fertility and Female Labor Supply in Latin America: New Causal Evidence". *Labour Economics* 14(3), 565–573.

Ebenstein, A. (2009). "When Is the Local Average Treatment Close to the Average? Evidence from Fertility and Labor Supply". *The Journal of Human Resources* 44(4), 955–975.

Frenette, M. (2011). "How Does the Stork Delegate Work? Childbearing and the Gender Division of Paid and Unpaid Labour". *Journal of Population Economics* 24(3), 895–910.

Goujon, M. (2008). "L'indice de développement humain : une évaluation pour la réunion". *Région et Développement* (27-2008).

Hirvonen, L. (2010). "The Effect of Children on Earnings Using Exogenous Variations in Family Size: Swedish Evidence". Working paper.

Hupkau, C. and M. Leturcq (2016). "Fertility and Labor Supply: New Evidence from the UK". Document de travail 220, INED.

Iacovou, M. (2001). "Fertility and female labour supply". ISER Working Paper Series 2001-19, Institute for Social and Economic Research.

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

Lewbel, A., Y. Dong, and T. T. Yang (2012). "Comparing features of convenient estimators for binary choice models with endogenous regressors". *Canadian Journal of Economics/Revue canadienne d'économique* 45(3), 809–829.

Li, H., J. Yi, and J. Zhang (2015). "Fertility, Household Structure, and Parental Labor Supply: Evidence from Rural China". Discussion Paper Series 9342, IZA.

Lopez de Lerida, J. (2005). "The Impact of Exogenous Variation in Family Size on Women's Labour Force Participation". Mimeo.

Lundborg, P., E. Plug, and A. W. Rasmussen (2014). "Fertility Effects on Female Labor Supply: IV Evidence from IVF Treatments". Discussion Paper Series 8609, IZA.

Maurin, E. and J. Moschion (2009). "The Social Multiplier and Labor Market Participation of Mothers". *American Economic Journal: Applied Economics* 1(1), 251–72.

Moschion, J. (2009). "Offre de travail des mères en France : l'effet causal du passage de deux à trois enfants". *Economie et Statistique* (422), 51–78.

Moschion, J. (2010). "Reconciling Work and Family Life: The Effect of the French Paid Parental Leave". *Annals of Economics and Statistics* (99/100), 217–246.

Rosenzweig, M. R. and K. I. Wolpin (1980). "Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models". *Journal of Political Economy* 88(2), 328–348.

Stock, J. and M. Yogo (2005). "Testing for Weak Instruments in Linear IV Regression", pp. 80–108. New York: Cambridge University Press.

Waldfogel, J. (1998). "Understanding the "Family Gap" in Pay for Women with Children". *The Journal of Economic Perspectives* 12(1), 137–156.