

## Appendix 2

### Understanding of Inheritance Rules from Gujarat Data

Table A2.1 Understanding of Inheritance Rule

	Freq.	Percent
Equal Shares to all sibs	152	14.66
Equal Shares to brothers	852	82.16
Oldest Brother gets more	2	0.19
No Standard Rule	19	1.83
Other / Don't know	12	1.16
Total	1,037	100

Table A2.2 Do Some Brothers Inherit Better Land than Others?

	Freq.	Percent
Yes	302	29.12
No	733	70.68
Other / Don't know	2	0.2
Total	1,037	100

Table A2.3 If one brother has more education or a better job, will he inherit less land?

	Freq.	Percent
Yes	70	6.75
No	965	93.06
Other / Don't know	2	0.2
Total	1,037	100

Source: Collected by author

Notes: These tables present summary statistics on perceptions of land inheritance rules and adherence to them. The data was collected by the author and is from a random sample of 1,037 respondents engaged in agriculture in Gujarat, India.

## Appendix 3

### Wage Gaps By Occupation in Rural India (within district OLS estimates)

	Occupational Wage Gaps by Farmer Landholdings		
	(1)	(2)	(3)
Comparison Group	All Farmers	<= 3 acres	> 3 acres
<i>A. Primary Occupation: Non-Farm Business</i>			
Wage Gap	-1.219 (1.603)	18.885*** (1.827)	-10.669*** (1.834)
N	30707	13621	22266
<i>B. Primary Occupation: Salaried Job</i>			
Wage Gap	21.610*** (2.051)	37.265*** (2.347)	11.290*** (2.258)
N	31637	14551	23196
<i>C. Primary Occupation: Non-Agricultural Labor</i>			
Wage Gap	-15.062*** (1.113)	4.695*** (1.236)	-26.601*** (1.398)
N	35019	17933	26578
Daily Wage for Comparison Farmers	81.902	50.933	97.171
Sex FE	Y	Y	Y
Age FE	Y	Y	Y
Education FE	Y	Y	Y
District FE	Y	Y	Y

Source: Indian Human Development Survey

Notes: This table computes differences in the average daily wage within districts by occupation in rural India. The sample is restricted to all individuals in rural India in the Indian Human Development Survey. Column 1 in Panel A reports the estimated OLS coefficient from a regression of the daily wage on a dummy variable coded as 1 if the main source of income is business and 0 if it is farming. Column 2 reports the same coefficient but restricting the comparison group to farmers with less than or equal to 3 acres of land. Column 3 restricts the comparison group to farmers with more than 3 acres of land. Panel B does similarly where the main occupation is instead a salaried job, and Panel C considers Non-agricultural wage work. All specifications include district fixed effects, non-parametric controls for sex, age (15 dummies, 0-80, 5 year intervals) and years of education (15 dummies, 0-15 years, 1 year intervals). Primary Source of Income defined as source of income with highest proportion relative to total income for an individual. A 'farmer' is defined as an individual whose highest proportion of income is from own agricultural cultivation or agricultural labor. The daily wage for farming is calculated as total farm profit divided by the number of days spent in agricultural labor, or the agricultural wage income in the case of agricultural labor.

## Appendix 4

### The Effects of Birth Order on Land Ownership (Sibling-Level Data)

Appendix 4 reports the coefficients on the birth order dummies from the following within-family regression:

$$Y_{ij} = \alpha_j + \gamma_z \sum_{z=1}^6 I(\text{Birthorder}_{ij} = z) + \mu_1 \text{Age\_Dummies}_{ij} + \mu_2 \text{Education\_Dummies}_{ij} + \eta_{ij}$$

Where  $Y$  is a dummy coded as 1 if current landholdings are greater than inherited landholdings and 0 otherwise, for sibling  $i$  of head  $j$ , and  $\alpha_j$  is a family fixed effect. Column (2) shows that latter-borns are less likely to experience an increase in their current landholdings over their inherited landholdings. In the main sample, 26% of household heads report that they experienced an increase in their landholdings over the prior two decades. Of these respondents, nearly 40% report receiving ‘gifts’ of land, a category distinct to inheriting, leasing or purchasing land. While no further details are given about these gifts in the 1999 wave, in the 2006 wave of the REDS survey, 80% of land leased in is from family members. The majority of these contracts are oral rather than written, they do not involve a fee and have no specified term. Taken together, these facts support the interpretation that latter-born siblings, unbound by social obligations, ‘lease’ their land to first-born siblings, and are more likely to specialise in non-agricultural occupations to the benefit of their family’s future consumption.

Dependent Variable	Inherited Land (Binary ) (1)	Land Increase (Binary) (2)	Change in Landholdings (Acres) (3)
2nd Born	-0.009 (0.008)	-0.016* (0.008)	-0.052 (0.051)
3rd Born	0.007 (0.011)	-0.032*** (0.011)	-0.083 (0.069)
4th Born	0.010 (0.014)	-0.049*** (0.015)	-0.269*** (0.092)
5th Born +	-0.001 (0.018)	-0.051*** (0.019)	-0.321*** (0.118)
Constant	0.542*** (0.112)	0.275** (0.123)	1.382* (0.751)
Family FE	Y	Y	Y
Age FE	Y	Y	Y
Depvar Mean	0.674	0.256	0.731
N	14773	14773	14773

Source: ARIS-REDS Dataset

Notes: This table reports within family estimates of the effect of birth order on the probability of inheriting land and changes in landholdings over time. The sample is restricted to all male siblings who reached the age of 10 years prior to death. In each family, one of the brothers is a household head in the main analysis. The data is at the sibling-level. Note, this data is reported for all siblings not just siblings residing in the household at the time of the survey. The dependent variable in Col 1-3 is the total number of man days spent in agriculture during the prior season. Columns 1-5 report the coefficient on a dummy for being the 2nd born sibling, 3rd born sibling, 4th born sibling and the 5th born or later sibling. The dependent variable in column 1 is a dummy variable coded as 1 if the sibling inherited land. The dependent variable in column 2 is a dummy coded as 1 if the sibling's current landholdings are greater than his inherited landholdings, and in column 3 it is the difference between current and inherited landholdings in acres. All specifications include family fixed effects, age fixed effects (0-100 years, 5 year intervals, 19 dummies), and education fixed effects (0-14 years, 1 year intervals, 13 dummies). The excluded group are first born siblings between the ages of 0-5 with less than an year of education. Robust standard errors are given in parentheses.

## Appendix 5

### The Effects of Inherited Land on Non-farm Business Ownership, Salaried Work and Non-agricultural Wage Work

Dependent Variable	Non-Farm Business Binary Variable		Salaried Job Binary Variable		Non-Agricultural Wage Work Log(Rs.)	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Inherited Land (Acres)	-0.001* (0.001)	- 0.007*** (0.002)	-0.001 (0.001)	-0.000 (0.003)	- 0.002*** (0.000)	-0.004 (0.003)
Mean of Dep. Var.	0.054	0.054	0.095	0.095	0.062	0.062
First Stage F-statistic	-	125.952	-	125.952	-	125.952
N	4809	4809	4809	4809	4809	4809
No. of Siblings FE	Y	Y	Y	Y	Y	Y
Parents Land FE	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y

Source: ARIS-REDS Dataset

Notes: This table reports estimates of the long-term effect of inherited land on non-farm business ownership, holding a salaried position and non-agricultural wage work. Columns 1, 3 and 5 report OLS coefficient estimates while columns 2, 4 and 6 report 2SLS estimates. The sample are Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. The data is at the household head-level. The dependent variable in all columns correspond to the primary status reported by the respondent in the REDS survey. In column 1 and 2 if the primary status is non-farm business then it is coded as 1, in cols 3 and 4 if it is a salaried position and if cols 5 and 6 if it is non-agricultural wage work. The instrument specification used is Predicted Share =  $1 / (1 + \text{Brothers})$ . The F-stat reported is the partial F-statistic for the instrument(s) (Cragg-Donald Wald F-statistic). Robust standard errors are given in parentheses. Standard errors are clustered at the family level for sibling-level regressions.

## Appendix 6

## Reduced Form Estimates by Birth Order (landless parents)

Dependent Variable	Non-Ag Occupation Binary (1)	Ag Labor Total Person days (2)	Land Improvement Log(Rs.) (3)	Household Consumption Log(Rs.) (4)	Total Value of Loans Log(Rs.) (5)	Rural-Urban Migration Binary (6)
First Born	0.029 (0.059)	-3.496 (2.239)	-0.028 (0.131)	-0.040 (0.067)	-0.273 (0.378)	0.001 (0.013)
Predicted Share*First Born	-0.073 (0.095)	2.334 (2.920)	-0.111 (0.212)	0.063 (0.101)	0.347 (0.594)	-0.033* (0.018)
Predicted Share*Latter Born	0.048 (0.083)	-3.764 (3.654)	-0.300** (0.146)	-0.030 (0.102)	0.242 (0.691)	-0.048*** (0.018)
Mean of Dep. Var.	0.551	5.837	0.192	10.139	1.138	0.017
N	1315	1315	1315	1315	1315	3899
No. of Siblings FE	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y
Parents Land FE	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y
Dowry, Education and Other Controls	Y	Y	Y	Y	Y	Y
Reject Equality of Coefs	N	N	N	N	N	N
Data	Head-Level	Head-Level	Head-Level	Head-Level	Head-Level	Sibling-Level

Source: ARIS-REDS Dataset

Notes: This table reports the reduced form regressions on a dummy for a First Born child, Predicted Share interacted with First Born, and Predicted Share interacted with Latter Born (a dummy for all household heads who are not First Borns) and is restricted to household heads whose parents did not own land. 'Dowry, Education and other controls' include own education, mother's education, father's education, average spacing between siblings (years), net dowry receipts (adjusted for inflation), birth order and age. Non-Ag occupation (col 1) is defined as the primary status reported by the respondent in the REDS survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. 'Ag labor' (col 2) is the total number of days of agricultural labor performed by the head in the prior season, Land improvement (col 3) is the log of the total expenditure on improvements in land (e.g. terracing, bunding, fencing, leveling, reclamation etc...) undertaken in the last 10 years. The dependent variable in col 4 is the natural logarithm of yearly household consumption. This includes food and non-food items and values home production at village-specific market prices. The dependent variable in col 5 is the log of the total value of all loans taken out in the last 5 years. The dependent variable in col 6 is a dummy variable for whether or not the sibling migrated to an urban area within the district or outside it. 'Reject Equality of Coefs' tests whether one can reject equality of coefficients for first-born and latter-born with a p-value less than 0.1. Robust standard errors are given in parentheses. Standard errors are clustered at the family for sibling-level regressions.

Appendix Table 7

## Reduced Form Estimates for Matrilineal and Patrilineal Communities

Dependent Variable Sample	Inherited Land (Acres) (1)	Non-Ag Occupation Main Sample, ARIS-REDS (Binary) (2)	HH Consumption Log(Rs.) (3)	Rural-Urban Migration (Binary) (4)	Land Owned (Acres) (5)	Non-Ag Occupation (Binary) (6)	HH Consumption Log (Rs.) (7)
Predicted Share* Patrilineal	7.50*** (0.670)	-0.112*** (0.035)	0.189*** (0.045)	-0.012*** (0.004)	1.620*** (0.254)	-0.056** (0.023)	0.143*** (0.031)
Predicted Share*Matrilineal	1.461** (0.625)	-0.1894 (0.102)	-0.084691 (0.104)	-0.035* (0.018)	0.429 (0.373)	0.074 (0.068)	0.010 (0.080)
Mean of Dep. Var.	4.120	0.298	10.441	0.011	1.652	0.430	13.413
N	4809	4809	4809	14773	11181	11181	11181
District FE	Y	Y	Y	Y	Y	Y	Y
No. of Siblings FE	Y	Y	Y	Y	Y	Y	Y
Parents Land FE	Y	Y	Y	Y	N	N	N
Reject Equality of Coefs	Y	N	Y	N	Y	Y	Y
Data Level	Head- Level	Head- Level	Head-Level	Sibling- Level	Head- Level	Head- Level	Head-Level

Source: ARIS-REDS Dataset

Notes: Columns 1-4 report reduced form estimates using Predicted Share as the instrument for Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. Matrilineal is a dummy that switches on for households in Kerala who belong to the Nair and Ezhava castes. Columns 5-7 the data used are Hindu male household heads in the 2004-2005 wave of the Indian Human Development Survey, who reside in rural areas. Matrilineal is a dummy that switches on for households in Kerala, Assam, Arunachal Pradesh, Meghalaya, Tripura or Nagaland. These are areas reported as having Matrilineal or Bilineal (referred to as 'Bilateral by Agarwal (2004)) Inheritance laws in Agarwal (2004). Non-Ag occupation is defined by the primary status reported by the respondent in the survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. Yearly household consumption includes food and non-food items, and values home production at village-specific market prices. Robust standard errors are given in parentheses. 'Reject Equality of Coefs' tests whether one can reject equality of coefficients between patrilineal and matrilineal interactions with predicted share with a p-value less than or equal to 0.1 Robust standard errors are given in parentheses. Standard errors are clustered at the family level for sibling-level regressions.

## Appendix 8

## Correlation Between Predicted Inheritance Shares and Observable Characteristics By Number of Siblings

<i>Dependent Variable</i>	Mean/ S.D.		Coefficient on Predicted Share			
	(1)	(2)	(3)	(4)	(5)	(6)
Age of Head	49.263 14.149	-1.138 (2.460)	6.058** (2.742)	-3.261 (3.198)	3.191 (3.740)	0.947 (1.085)
Father's Education (Years)	1.428 2.759	0.202 (0.421)	-0.860** (0.414)	0.567 (0.671)	0.063 (0.788)	0.103 (0.205)
Mother's Education (Years)	0.407 1.440	0.198 (0.187)	-0.314* (0.163)	0.610* (0.319)	0.343 (0.455)	0.146 (0.102)
Father in Agriculture (Primary Occupation)	0.866 0.340	0.021 (0.062)	0.027 (0.067)	0.105 (0.096)	-0.017 (0.082)	0.008 (0.027)
Dowry Received Log(Rs. +1)	4.929 3.977	0.159 (0.578)	-0.073 (0.631)	-0.726 (0.828)	-1.703** (0.794)	-0.377 (0.261)
Age when Headship Assumed (Years)	33.397 9.988	-1.939 (1.786)	0.804 (1.902)	-6.436** (2.736)	-3.547 (2.725)	-1.450* (0.790)
Age of Marriage (Years)	22.193 5.173	-1.103 (0.761)	0.693 (0.797)	-0.184 (1.042)	-0.575 (1.083)	-0.367 (0.347)
Birth Order	2.585 1.796	-0.045 (0.153)	-0.078 (0.225)	-0.335 (0.395)	0.140 (0.475)	-0.112 (0.107)
Sibling Spacing (Years)	3.987 2.810	-1.155** (0.586)	-0.794 (0.491)	-0.445 (0.544)	-0.751** (0.382)	-0.579** (0.230)
N	4809	729	811	777	726	4809
Sample	All	2 siblings	3 siblings	4 siblings	5 siblings	All
No. of Siblings FE	-	N	N	N	N	Y
Parents Land FE	-	Y	Y	Y	Y	Y
District FE	-	Y	Y	Y	Y	Y

Data Source: ARIS-REDS Dataset

Notes: This table presents summary statistics (mean and standard deviation) in Column 1 and assesses the conditional independence assumption of the instrument in Col 2-6, by seeing if it is independent of a number of characteristics of the household head and the head's parents. The sample is restricted to Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. The data is at the household head level. †Columns 2-6 report the coefficient estimate on the instrument, Predicted Share =  $1/(1+\text{Brothers})$ , from a reduced form regression of the dependent variable on the instrument. Columns 2-5 assesses balance for household heads with varying numbers of siblings, while Column 6 includes all household heads. 'Father in agriculture' is coded as 1 if the primary occupation of the head's father is agriculture. 'Dowry received' reports the natural logarithm of the value of dowry payments given to the head or his parents at the time of marriage. 'Age when Headship Assumed' reports the age at which the respondent assumed headship of the household. 'Birth Order' is an integer value that is rising in parity (1 if eldest) and 'Sibling Spacing' computes the average interval between sibling births in number years and is winsorized (1%). All specifications include district fixed effects (99 dummies), fixed effects for family landholdings (0-80+ acres, 5 acre intervals, 15 dummies) and the number of siblings (14 dummies). The excluded group are heads who are only children, from West Godavari district in Andhra Pradesh with family landholdings between 0-5 acres. Brothers are defined as male siblings who grew up to at least the age of 10. Results are robust to alternative definitions and using ever born siblings. Robust standard errors are given in parentheses.

Appendix 9  
Robustness of 2SLS Estimates to Alternative Instrument Specifications

Instrument Specification	Linear (No. of Brothers)	Predicted Share	Predicted Land	Log(Predicted Land)	Non-Parametric (Brother Dummies)
	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Dependent Variable: Non-Ag Occupation (Binary)</i>					
Inherited Land (Acres)	-0.020*** (0.005)	-0.018*** (0.005)	-0.002 (0.002)	-0.042*** (0.007)	-0.017*** (0.005)
Mean of Dep. Var.	0.298	0.298	0.298	0.298	0.298
First Stage F-Statistic	128.137	125.952	101.458	141.607	22.433
N	4809	4809	4809	4809	4809
<i>Panel B. Dependent Variable: Yearly Household Consumption, Log (Rs.)</i>					
Inherited Land (Acres)	0.032*** (0.006)	0.027*** (0.006)	0.022*** (0.003)	0.052*** (0.007)	0.031*** (0.006)
Mean of Dep. Var.	10.442	10.442	10.442	10.442	10.442
First Stage F-Statistic	128.137	125.952	101.458	141.607	22.433
N	4809	4809	4809	4809	4809
No. of Siblings FE	Y	Y	Y	Y	Y
Parents Land FE	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y

Data Source: ARIS-REDS Dataset

Notes: This table tests the robustness of 2SLS estimates of the long-term effect of inherited land on occupational choice and household consumption to alternative specifications of the instrument. The sample is restricted to Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. The data is at the household head level. Column 1 reports 2SLS estimates with the instrument specified as the (linear) number of brothers, in column 2 it is 'Predicted Share' =  $1/(1+\text{Brothers})$ , in column 3 it is 'Predicted Land' =  $\text{Family Land}/(1+\text{Brothers})$ , in column 4 it is  $\text{Log}(\text{Predicted Land})$ , and in column 5 it is the a set of dummies for the number of brothers (8 dummies in total, I report the coefficients for up to 5 brothers which account for 98.11% of sample). The dependent variable in Panel A is Non-Ag occupation and is defined by the primary status reported by the respondent in the REDS survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. The dependent variable Panel B is the natural logarithm of yearly household consumption. This includes food and non-food items, and values home production at village-specific market prices. All specifications include district fixed effects (99 dummies), fixed effects for family landholdings (0-80+ acres, 5 acre intervals, 15 dummies) and the number of siblings (14 dummies). The excluded group are heads who are only children from West Godavari district in Andhra Pradesh with family landholdings between 0-5 acres. The F-stat reported is the partial F-statistic for the instrument(s) (Cragg-Donald Wald F-statistic). Robust standard errors are given in parentheses.

## Appendix 10

### Reduced Form Simulations of Selective Migration

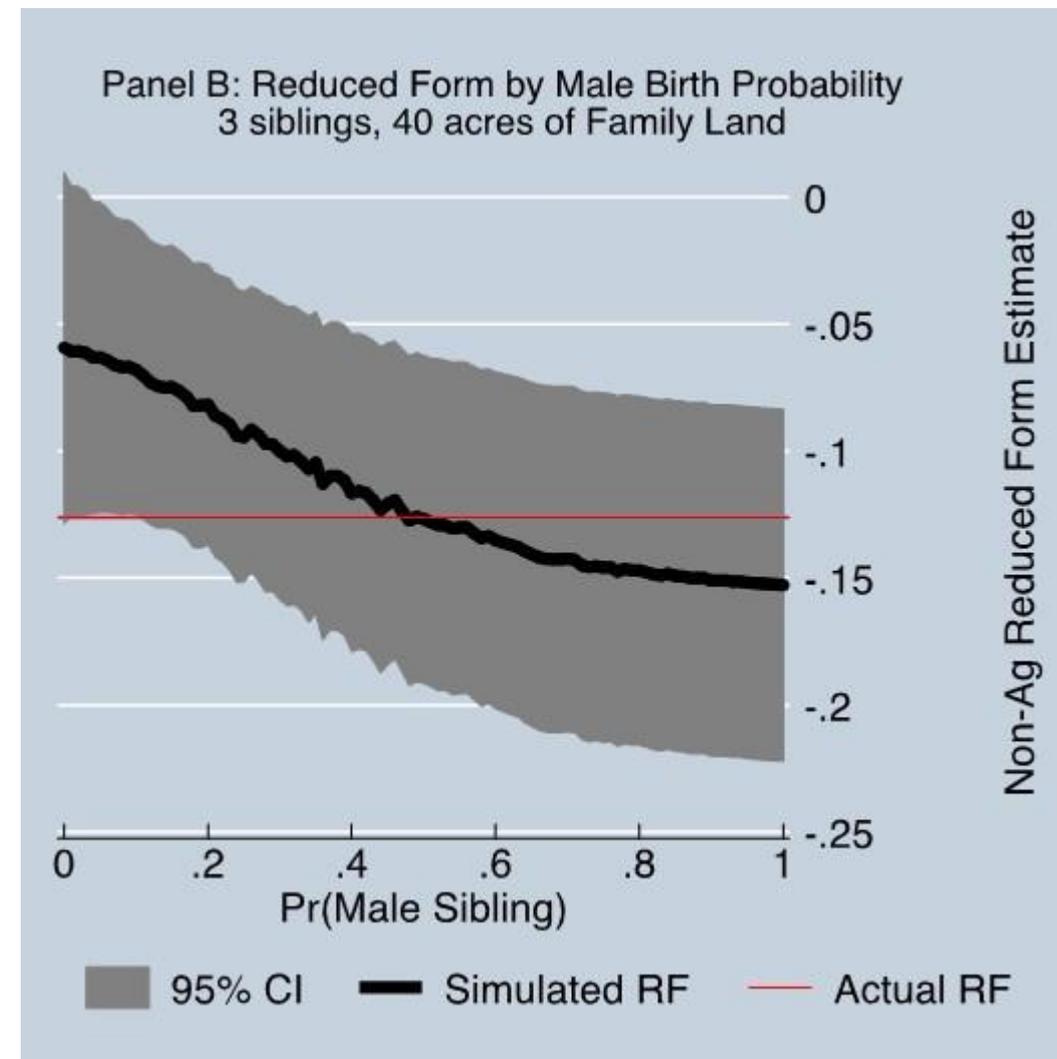
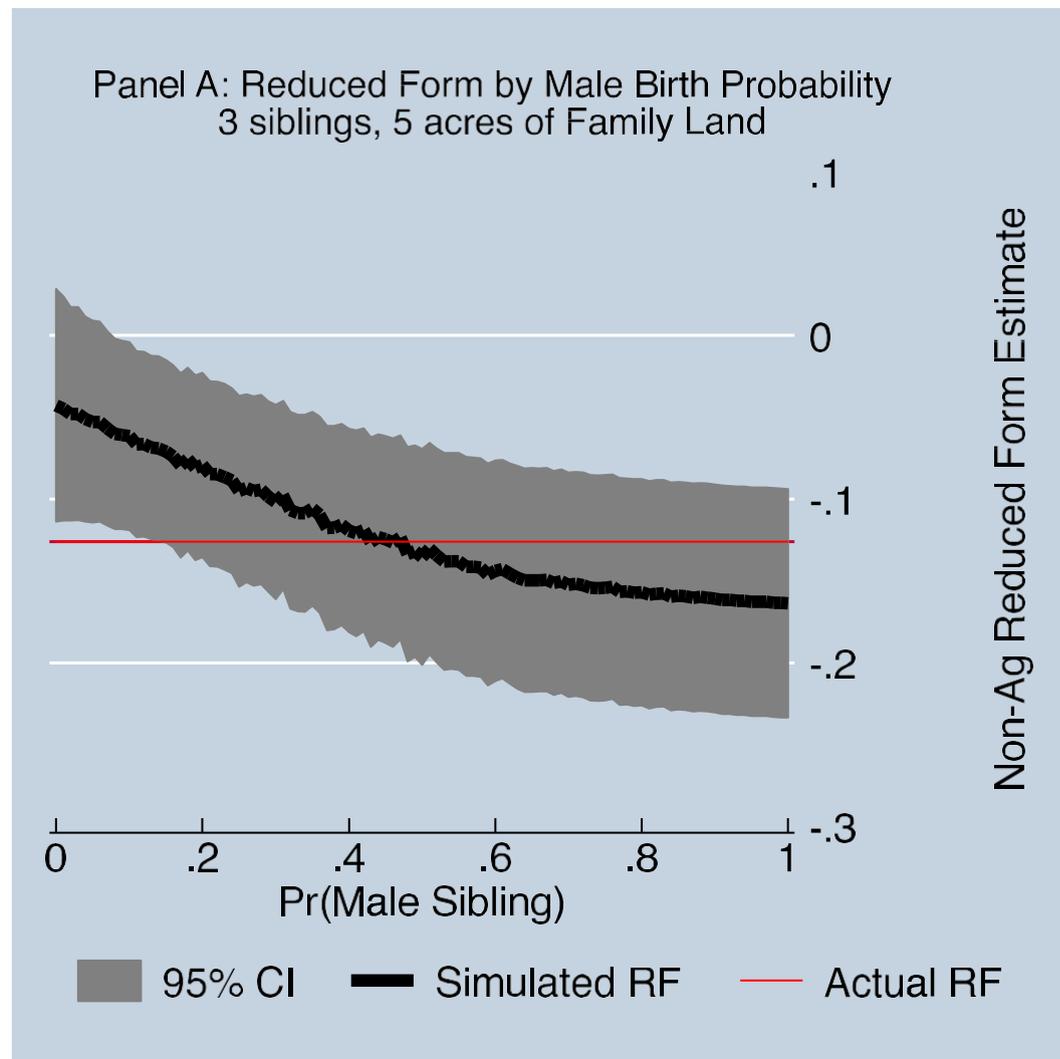
A potential selection concern arises from the fact that all surveyed households in the REDS dataset reside in rural areas. As such, a subset of individuals who migrate to urban areas after inheriting land are not sampled and may compromise the estimated relationships. For example, if these ‘missing migrants’ inherited small amounts of land and subsequently took up non-agricultural jobs in urban areas, I would underestimate the negative effect of land. Conversely, if the migrants inherited large amounts of land and then took up a non-agricultural occupation, I would overestimate the negative effect of land. Given the nature of the REDS data these migrants would need to result from the movement of entire *families* to urban areas; household heads report the location of their siblings irrespective of where they reside. While nationally representative estimates of the extent of permanent rural-urban *family* migration are not available, studies suggest this form of migration is extremely rare in India ((Munshi and Rosenzweig 2007)). In the REDS sibling data, just 1.1% of 16,130 male siblings have migrated to urban areas. Foster and Rosenzweig (2007) estimate the *individual* rural to urban migration rate for males aged 15-24 for each decade between 1961 and 2001 using the corresponding Indian censuses. They find that migration rates vary from 3% to 5% for each of the decades between 1961-2001, suggesting very limited migration even when considering the movement of individuals rather than entire families.

The estimated negative relationships between inherited land and both migration and entering non-agricultural work suggest that migrants would need to have large landholdings in order to overturn the estimates. It is worth noting that this is a hypothetical at odds with the estimated negative causal effect of inherited land on urban migration. Additionally, these landholdings cannot be so large that they have little influence on the 2SLS estimates. The latter restriction is a result of the nonlinearity of the estimated relationship between inherited land and both occupational choice and migration as suggested by Panel B of Table 3 and Panel A & B of Figure 4 and the weighting structure of 2SLS with covariates. The estimated slopes are especially negative and precise for inheritances of up to 4 acres of land. However, for those inheriting more than 4 acres, the qualitative nature of the relationship is unclear and the estimates are imprecise. As such, migrants with very large landholdings would be included in covariate-specific LATE’s (i.e. the 2SLS estimate computed for subsets of the sample covariates) that are qualitatively different in sign from the overall LATE and have little variation in the instrument as a consequence of having few observations and are therefore not heavily weighted in the overall LATE.

These restrictions suggest that migrants whose parents owned intermediate amounts of land – i.e. covariate values that occur frequently in the data and drive the negative estimated relationships – would be the most likely to overturn the reduced form estimates. Having specified parents’ landholdings, the sibling sex composition of migrants would determine their inherited landholdings. Appendix C3 models the sibling sex composition of migrants as resulting from a series of draws from a binomial distribution and estimates the reduced form for occupational choice – all urban migrants are assumed to hold non-agricultural occupations – under varying probabilities of success (i.e. the probability of drawing a male sibling). These simulations quantify how skewed the sibling sex composition of migrants would need to be to overturn the reduced form estimate.

The simulations add observations to mimic a 10% rural to urban migration rate: the census-based *individual* urban migration rate for the

three decades preceding the REDS survey. In both panels the reduced form estimate from the main specification is indicated by the horizontal red line, while the grey area shows the 95% confidence interval for the simulated reduced form coefficients. Panel A shows the estimated reduced form coefficients when migrants are assumed to have the most frequent sibling and parents' land combination: 3 siblings and parents who own 5 acres of land, while in Panel B they are assumed to have parents who own 40 acres land (95th percentile for parents' landholdings). As suggested by the discussion above, migrants whose parents have smaller landholdings (Panel A) have a greater influence on the reduced form estimates than those with large landholdings (Panel B). However, even in Panel B the probability of a male sibling occurring would need to be less than 0.17 – the point at which the red line leaves the confidence interval – in order to overturn the reduced form relationship. This is substantially lower than the biological probability of a male and the observed ratio of brothers to siblings for rural to urban migrants in the IHDS data: 0.53.



#### Appendix 10

#### Simulations of Selective Migration and the Reduced Form Effect on Occupational Choice

Source: ARIS-REDS Dataset

Notes: These figures plot the results of simulations intended to test the robustness of the reduced form relationship for non-agricultural occupation to selective family migration. Both panels add the simulated data to the main sample of Hindu male household heads whose parents owned land in the 1999 wave of the ARIS-REDS dataset. Both panels assume a family migration rate of 10%. Panel A assumes the migrants have 3 siblings and their parents own 5 acres of land, while Panel B assumes that the migrants have 3 siblings and their parents own 40 acres of family land. Where  $N$  is the number of siblings assigned to the migrants, the program takes  $N$  draws from a binomial distribution with a success  $K$  for each of the migrants.  $K$  is varied from 0 to 1 in intervals of 0.01 and the reduced form relationship is estimated for each of these values and plotted with the 95% confidence interval using robust standard errors. The red line in each panel shows the estimated reduced form estimate with the main specification: -0.126. In each case the missing migrants are assumed to come from Allahabad District in Uttar Pradesh. All specifications include district fixed effects (99 dummies), fixed effects for parents landholdings (0-80+ acres, 5 acre intervals, 15 dummies) and the number of siblings (14 dummies). The excluded group are heads who are only children, from West Godavari district in Andhra Pradesh with family landholdings between 0-5 acres.

## Appendix 11

## The Effect of Inherited Land on By Sibling Cohort Size (2sls Estimates)

<i>Dependent Variable</i>	2SLS Estimates					
	Mean/ S.D.	Effect of Inherited Land on Outcomes by Sibling Cohort Size†				
	Full Sample	2 siblings	3 siblings	4 siblings	5 siblings	Full Sample
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Dependent Variable: Non-Agricultural Occupation</i>						
Inherited Land (Acres)	0.298 (0.458)	-0.007 (0.013)	-0.043*** (0.014)	-0.010 (0.012)	-0.021* (0.011)	-0.018*** (0.005)
N	4809	729	811	777	726	4809
<i>Panel B. Dependent Variable: Log(Household Consumption)</i>						
Inherited Land (Acres)	10.442 (0.614)	0.032** (0.016)	0.005 (0.017)	0.006 (0.014)	0.053*** (0.011)	0.027*** (0.006)
N	4809	729	811	777	726	4809
No. of Siblings FE	-	N	N	N	N	Y
Parents Land FE	-	Y	Y	Y	Y	Y
District FE	-	Y	Y	Y	Y	Y

Source: ARIS-REDS Dataset

Notes: This table presents 2SLS estimates for the effect of inherited land on occupational choice and household consumption by sibling cohort size. The sample is restricted to Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. The data is at the household head level. †Columns 2-6 report the 2SLS coefficient on inherited landholdings. Columns 2-5 report the coefficient for household heads with varying numbers of siblings, while Column 6 includes all household heads. Non-Ag occupation is defined by the primary status reported by the respondent in the REDS survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. Log(Household Consumption) is the natural logarithm of yearly household consumption which includes food and non-food items, and values home production at village-specific market prices. Brothers are defined as male siblings who grew up to at least the age of 10. Results are robust to alternative definitions and using ever born siblings. Robust standard errors are given in parentheses.

## Appendix 12

## Robustness of 2sls Estimates to Sex Selective Fertility Preferences

Specification	Main	Exact Permuation	Sibling Spacing Controls	Youngest Sibling is Female
	(1)	(2)	(3)	(4)
<i>Panel A. Dependent Variable: Non-Ag Occupation (Binary)</i>				
Inherited Land (Acres)	-0.018*** (0.005)	-0.029*** (0.007)	-0.018*** (0.005)	-0.018*** (0.006)
Mean of Dep. Var.	0.298	0.298	0.298	0.292
N	4809	4809	4809	1947
<i>Panel B. Dependent Variable: Log(Household Consumption)</i>				
Inherited Land (Acres)	0.027*** (0.006)	0.016** (0.008)	0.027*** (0.006)	0.035*** (0.008)
Mean of Dep. Var.	10.442	10.442	10.442	10.455
F-stat (First Stage)	125.952	89.579	132.955	59.106
N	4809	4809	4809	1947
No. of Siblings FE	Y	Y	Y	Y
Parents Land FE	Y	Y	Y	Y
District FE	Y	Y	Y	Y
Sibling Sex Permuation FE	N	Y	N	N
Sibling Spacing FE	N	N	Y	N
Instruments	Pred Share	Pred Share	Pred Share	Pred Share
Sample	All	All	All	Last Born Sis

Source: ARIS-REDS Dataset

Notes: This table tests the robustness of 2SLS estimates of the long-term effect of inherited land on occupational choice and household consumption to sex selection and son-preferring differential stopping behavior, both of which stand to violate the conditional independence assumption of the instrument. The sample is restricted to Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. The data is at the household head level. Column 1 reports 2SLS estimates for the main specification used in Table 3. Column 2 includes fixed effects for the exact permutation of the sex of siblings born prior to the head of the household (i.e. MMF, FFM etc..) and includes 223 dummy variables. Column 3 includes fixed effects for the average birth spacing (in years) between siblings (0-10+ years, 6 month intervals, 18 dummies). Column 4 limits the sample to the subset of heads whose youngest sibling is female. This is under the assumption that those families who stop on a girl are more likely to have satisfied a resource constraint than stopped because of son-preferring differential stopping behavior. The dependent variable in Panel A is Non-Ag occupation and is defined by the primary status reported by the respondent in the REDS survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. The dependent variable in Panel B is the natural logarithm of yearly household consumption. This includes food and non-food items, and values home production at village-specific market prices. The F-stat reported is the partial F-statistic for the instrument(s) (Cragg-Donald Wald F-statistic). Robust standard errors are given in parentheses.

## Appendix 13

## Robustness of 2SLS Estimates To Controls For Parent's Dowry Expenditure and Head's Education

Dependent Variable	First Stage		Main	Education Controls	Dowry Controls	Main	Education Controls	Dowry Controls
	Education (Years)	Net Dowry (Binary)	Non-Ag Occupation (Binary)	Non-Ag Occupation (Binary)	Non-Ag Occupation (Binary)	HH Consumption Log(Rs.)	HH Consumption Log(Rs.)	HH Consumption Log(Rs.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A : First Stage (Landed Only)</i>								
Predicted Share	1.397*** (0.386)	-0.074 (0.030)	-	-	-	-	-	-
Mean of Dep. Var.	5.847	0.708	-	-	-	-	-	-
F-stat (First Stage)	13.119	6.168						
N	4809	4809	-	-	-	-	-	-
<i>Panel B : First Stage (Landed and Landless)</i>								
Predicted Share	0.908 (0.583)	-0.069* (0.038)	-	-	-	-	-	-
Predicted Share*Landed	0.621 (0.579)	-0.010 (0.036)						
Mean of Dep. Var.	2.422	3.313	-	-	-	-	-	-
F-stat (First Stage)	5.540	0.711						
N	6124	6124	-	-	-	-	-	-
<i>Panel C: 2SLS Estimates (Landed Only)</i>								
Inherited Land (Acres)	-	-	-0.018*** (0.005)	-0.015*** (0.005)	-0.021*** (0.005)	0.027*** (0.006)	0.028*** (0.006)	0.024*** (0.006)
Mean of Dep. Var.	-	-	0.298	0.298	0.298	10.442	10.442	10.442
F-stat (First Stage)			125.952	124.544	118.234	125.952	118.234	124.544
N	-	-	4809	4809	4809	4809	4809	4809
No. of Siblings FE	Y	Y	Y	Y	Y	Y	Y	Y
Parents Land FE	Y	Y	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y	Y	Y
Education FE	N	N	N	Y	N	N	Y	N
Net Dowry FE	N	N	N	N	Y	N	N	Y

Data Source: ARIS-REDS Dataset.

Notes: This table tests the robustness of 2SLS estimates of the long-term effect of inherited land on occupational choice and household consumption to controls for net dowry receipts and education controls. The sample is restricted to Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. The data is at the household head level. Column 1 in Panel A reports the first stage for years of education using the 'predicted share' instrument. Column 2 in Panel A reports the first stage for net dowry. This variable is coded as 1 if the net dowry receipts of the household are above the median. The latter is calculated as the net sum of all dowry payments and receipts for the parents which are reported for each sibling of the head of the household adjusted for inflation using the CPI. 37% of the sample do not report paying or receiving dowry. Panel B tests whether the first stage for education and net dowry varies differentially for the landed and the landless. Columns 3-5 in Panel C report the 2SLS estimates of inherited land on non-agricultural occupation. Non-Ag occupation is defined by the primary status reported by the respondent in the REDS survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. Columns 6-8 in Panel B report the 2SLS estimate of inherited land on the log of household consumption. This includes food and non-food items, and values home production at village-specific market prices. Columns 4 and 7 include fixed effects for years of education of the household head, 12 dummies, 1 year intervals. Columns 5 and 8 include fixed effects for the net dowry receipts. This calculates the net difference between dowry received and spent by the head's parents for all siblings, and then creates 19 dummies (Rs. -50,000 - Rs. 50,000+, Rs. 5000 intervals). The F-stat reported is the partial F-statistic for the instrument(s) (Cragg-Donald Wald F-statistic). Robust standard errors are given in parentheses.

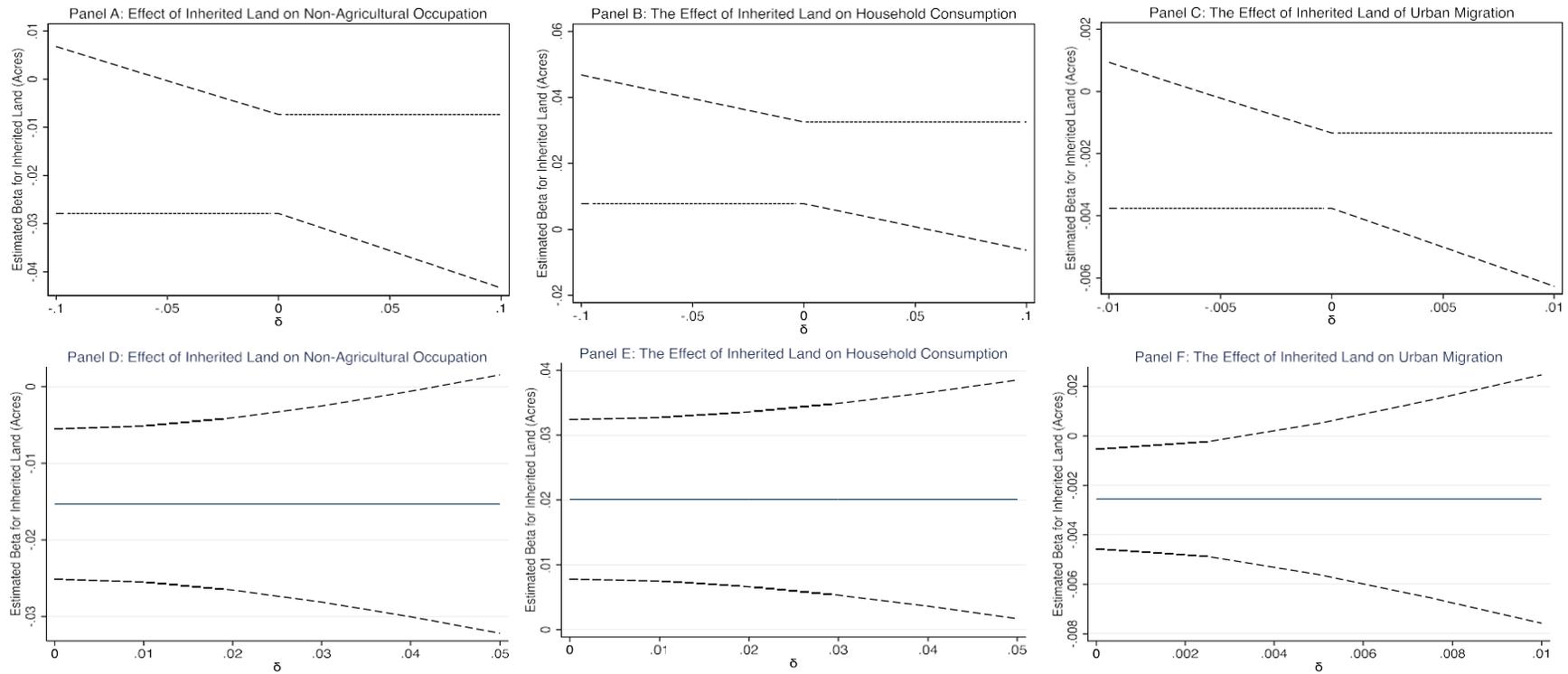
## Appendix 14

## Robustness of 2SLS Estimates to Alternative Definitions of Occupational Choice and Birth Orders Controls

Specification	Main	Definition Change	Main	Eldest Son Control	Main	Eldest Son Control
Dependent Variable	Non-Ag Occupation Binary (1)	Non-Ag Occupation Binary (2)	Non-Ag Occupation Binary (3)	Non-Ag Occupation Binary (4)	Household Consumption Log(Rs.) (5)	Household Consumption Log(Rs.) (6)
Land (Acres)	-0.018*** (0.005)	-0.016*** (0.004)	-0.004 (0.006)	-0.012 (0.007)	0.037*** (0.007)	0.055*** (0.014)
First Born	-	-	0.195*** (0.058)	0.173*** (0.059)	0.216*** (0.065)	0.296*** (0.084)
Land*First Born	-	-	-0.045*** (0.014)	-0.036** (0.015)	-0.033** (0.016)	-0.052** (0.022)
Mean of Dep. Var.	0.298	0.265	0.298	0.265	10.442	10.442
F-stat (First Stage)	128.137	125.952	26.667	26.667	26.667	29.852
N	4809	4809	4809	4809	4809	4809
Age FE	N	N	Y	Y	Y	Y
No. of Siblings FE	Y	Y	Y	Y	Y	Y
Parents Land FE	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y
Eldest Son Dummy	N	N	N	Y	N	Y
Non-Ag Definition	Primary Status	Majority Income	Primary Status	Primary Status	-	-
Eldest Son Dummy	No	No	No	Yes	No	Yes

Source: ARIS-REDS Dataset

Notes: This table tests the robustness of 2SLS estimates of the effect of inherited land on occupational choice to an alternative definition of occupational choice (Columns 1 and 2). This table also tests the robustness of heterogeneous effects by birth order to additional controls. The sample is restricted to Hindu male household heads whose parents owned land in the 1999 ARIS-REDS's survey. The data is at the household head level. Column 1 reports the 2SLS estimates using the main specification with the instrument as predicted share. In column 1, Non-Ag occupation is defined as the 'primary status' reported by the respondent in the REDS survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. Column 2 reports estimates using the definition of Non-Ag occupation as 'majority income'. This is defined as whether the majority of the respondent's income comes from activities not related to self-cultivation or agricultural labor using income data. Columns 3-6 report the 2SLS coefficients on inherited land, first-born - a dummy coded as 1 if the household head was the first-born child in his family - and their interaction. The two endogenous variables are instrumented with two instruments : Predicted Share = (1/1+Brothers) and the interaction between Predicted Share and First Born. These coefficients are from the main specification. Columns 4 and 6 include a dummy variable coded as 1 if the respondent is the eldest son but not the first born in the family. The dependent variable in columns 3 and 4 is Non-Ag occupation and is defined by the primary status reported by the respondent in the REDS survey. The variable is coded as 0 if this is self-cultivation or agricultural labor and 1 otherwise. The dependent variable columns 5 and 6 is the natural logarithm of yearly household consumption. This includes food and non-food items, and values home production at village-specific market prices. The F-statistic reported is the partial F-statistic for the instrument(s) (Cragg-Donald Wald F-statistic). Robust standard errors are given in parentheses.



## Appendix 15

### Sensitivity of Reduced Form Estimates to Violation of Exclusions Restriction

Source: ARIS-REDS Dataset

Notes: These figures show how large the exclusion restriction violation would need to be in order for the 2SLS estimates to be inconsistent. Panels A-C use the 'union of confidence intervals' approach, while Panels D-F use the local-to-zero approach that uses a normal distribution to model the prior for the size of the violation, where  $N(0, \delta^2)$ . (Conley et al. 2012). The dashed lines plot the 95% confidence intervals where  $\gamma$  is the size of the exclusion restriction violation such that  $\gamma \in [-2\delta, 2\delta]$ . Note, Panels A-C report non-symmetric confidence intervals. These figures were produced using the 'plausexog' code produced by Damian Clarke (2014).