How Love Conquered Marriage: Theory and Evidence on the Disappearance of Arranged Marriages

Gabriela Rubio*

February 2014

University of California, Merced

The most up-to-date version of this paper is available

at http://gabrielarubio.bol.ucla.edu/

Abstract

This paper documents the sharp and continuous decline of arranged marriages (AM) around the world during the past century, and describes the factors associated with this transition. To understand these patterns, I construct and empirically test a model of marital choices that assumes that AM serve as a form of informal insurance for parents and children, whereas other forms of marriage do not. Children accepting the AM will have access to insurance but might give up higher family income by constraining their geographic and social mobility. Children in love marriages (LM) are not geographically/socially constrained; they can look for the partner with higher labor market returns and have access to better remunerated occupations. The model predicts that AM disappear when the net benefits of the insurance arrangement decrease relative to the (unconstrained) returns outside of the social network. Using consumption and income panel data from the Indonesia Family Life Survey, I show that consumption of AM households does not vary with household income (while consumption of LM households does), consistent with the model's assumption that AM provides insurance. I then empirically test the main predictions of the model. I use the introduction of the Green Revolution (GR) in Indonesia as a quasi-experiment. First, I show that the GR increased the returns to schooling and lowered the variance of agricultural income. Then, I use a difference-in-difference identification strategy to show that cohorts exposed to the GR experienced a faster decline in AM as predicted by the theoretical framework. Second, I show the existence of increasing divorce rates among couples with AM as their insurance gains vanish. Finally, using the exogenous variation of the GR, I find that couples having an AM and exposed to the program were more likely to divorce, consistent with the hypothesis of declining relative gains of AM.

JEL: B52, C72, D13, J11, J12, O15, O53, O55.

*gabrielarubio@ucla.edu, grubio4@ucmerced.edu. I am truly indebted to my advisor, Adriana Lleras-Muney, for uncountable hours of discussion, guidance and support. I am also particularly grateful to Leah Boustan, Paola Giuliano, Maurizio Mazzocco and Robert Jensen for their encouragement, numerous discussions and guidance. And I am especially indebted to George Georgiadis for all his help and advice throughout all the stages of this project. I would also like to thank Kathleen McGarry, Maria Casanova, Dora Costa, Till von Wachter, Aprajit Mahajan, and Simon Board for helpful comments and suggestions. I have also benefited from participants at the Applied Micro Lunch Meeting, the Albert Family Proseminar at UCLA, the 2013 All-California Labor Economics Conference, the Applied Microeconomics Seminar at University of California at Riverside, and from seminars at Simon Fraser University, Eastern Connecticut State University, University of California at Merced, University of Illinois at Chicago and NERA. I would like to thank Esther Duflo for making available the data used in her paper "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." All remaining errors are mine.

1 Introduction

At the beginning of the 20th century, 72% (or more) of all marriages in Asia and Africa were arranged by the families of the couple. Throughout the last century, these marriages have decreased by approximately 40%. This paper uses a variety of sources to document this continuous and large decline in arranged marriages in several countries of Asia and Africa (Turkey, Saudi Arabia, Israel, Japan, Korea, China, Taiwan, Indonesia, Malaysia, Cambodia, Vietnam, Sri Lanka, Nepal, Togo and Ghana). Although a few countries (India, Pakistan and Bangladesh) do not follow the same patterns (arranged marriages still represent at least 95% of all marriages), young cohorts living in these countries' urban areas have also started to move away from arranged marriages. The goal of this paper is to understand the main driver(s) of the transition by proposing and testing empirically a model of marital choices. I first show that this transition away from arranged marriages in favor of self-choice or "love" marriages is correlated with increases in education, formal employment, urbanization, and declines in agriculture. These trends are common in all the countries where micro-data is available, suggesting that despite having different institutions at work, there is a fundamental economic explanation behind these changes in marriage institutions.

Based on these patterns, I build a simple model of marriage choice. I assume that arranged marriages serve as a form of informal insurance (as suggested in the literature of sociology, anthropology and economics, e.g. Rosenzweig and Stark (1989)) whereas other marriages (outside one's networks) do not. In the first period, parents spend resources on educating their offspring and looking for a spouse for an arranged marriage. Both investments have pay-offs in the second period when the child earns an education dependent wage subject to a shock, transfers a share back to her parents and decides to marry the parental arranged spouse or choose her own spouse. Arranged marriages provide insurance (to both parents and children) because parents (but not the children) can observe who within their network faces income shocks that are negatively correlated with their child's shocks (by observing the entire shock histories of other households and by having repeated interaction with them). However, arranged marriages come with a cost: they constrain the choice set and the mobility of the child, thereby reducing her potential income, compared to the case in which she has the option to move geographically, find a more lucrative occupation or a spouse with higher earnings. Thus, there is a trade-off between marrying within the network and marrying outside the network: individuals might be willing to give up potentially higher income in exchange for a risk-sharing agreement. The model predicts that arranged marriages disappear when the net benefits of the insurance arrangement decrease relative to the (unconstrained) returns outside of the social network. When this is the case, parents invest in more education for the child, effectively increasing her outside option and, thus, the probability that she will reject the arranged marriage.

In this framework, the decline in arranged marriages and the increase in years of schooling are endogenously determined; more specifically, the theoretical framework predicts that when the returns to schooling increase or the variance of income declines, the demand for insurance decreases — households pay a greater cost or "premium" for the insurance. I extend the base model in two directions that provide additional testable implications. First, motivated by the fact that divorce has a low cost in some of the countries studied (e.g. Indonesia), I introduce the possibility of divorce in a third period. Holding the cost of divorce constant, the model implies that couples in arranged marriages will have higher rates of divorce as their insurance gains vanish. Second, I assume two children within the household and theoretically explore how parents decide which child to offer an arranged marriage. I show that when social networks are small (equivalent to have insurance partners with positively correlated shocks), parents have incentives to arrange the marriage of only one child and they choose to marry only the child with the lowest expected return in the labor market.

The second part of the paper uses micro data to test the model. First, I test the main assumption of the model that arranged marriages provide insurance to the couple.[1] I implement a standard test of full insurance using the first three waves (1993, 1997 and 2000) of the Indonesia Family Life Survey (IFLS). This is the only data set containing both marriage type, and monthly consumption and income data of couples. For this test, I assume that the relevant unit for risk sharing is the village or town. If there is full insurance, the consumption of each individual household should not depend on its income but should have a one-to-one relationship with the average consumption of the insurance group. Although the test is formally rejected for both types of marriages, the results show that the consumption of arranged marriage couples does not depend on their income.[2] In contrast, the consumption of couples in love marriages varies significantly with their income, suggesting that individuals in arranged marriages are better insured. These results are consistent with Rosenzweig and Stark (1989), who show that households in India use the marriage of their children as insurance to mitigate the effect of profits shocks on consumption.

I then test the two main predictions of the model, namely that factors that lower the demand for informal insurance (increasing returns to schooling and decreasing income variance) accelerated the transition to love marriages and increased the investment in education. To test these predictions, I use the gradual introduction of the Green Revolution in Indonesia as a quasi-experiment. The Green Revolution refers to a series of technological innovations associated with the diffusion of higher yield variety seeds in developing countries in the late 1960s, which increased the returns to schooling and agricultural income (Foster and Rosenzweig, 1996).[3] I collected data from the 1963 and 1983 Indonesian agricultural census (and other printed sources) on the intensity of the Green Revolution by district (% of inputs use, % of land covered, rice yields, among others). I combine this information with the 1976 and 1995 census, the early 1980s socioeconomic surveys and the 1993 wave of IFLS to study the effect on labor market outcomes.

The results show that the Green Revolution increased returns to schooling by an additional 2.1% to 4.7% per additional year of schooling, it increased mean income of agricultural households, and importantly, it decreased their income variance by 8.1% and 8.3%, respectively. To identify how these changes affected the outcomes of interest (arranged marriages and education), I compare cohorts before and after the introduction of the program, in areas with high and low program intensity. As predicted, the Green Revolution resulted in a decline of 9 to 20 percentage points in the probability of having an arranged marriages for the cohort exposed to the Green Revolution, and in an increase in education of 0.3 to 0.5 years of schooling for the same cohort. The results are robust to adding controls for other programs implemented during the same period (school construction program, and expansion of water and sanitation supply), to adding controls at district level from the 1971 census, and to instrumenting the intensity of the Green Revolution with agricultural characteristics of 1963 (in order to mitigate concerns of potential endogeneity). In addition, they are also robust to using splines defined as the 4 quartiles of the treatment intensity distribution, to allowing for concave or convex effects by adding a quadratic term of the treatment variable, and to using an alternative

[1] I only have information on the marital arrangement, and income and consumption of children. In principle, their parents should also have access to insurance but I cannot test it with the current data.

[2] The results on aggregate consumption drive the rejection of the test for both samples.

[3] The Green Revolution in Indonesia was introduced from the mid-1960s to the mid-1980s and consisted of the expansion of higher yield rice seeds across regions in addition to availability of fertilizer and pesticides, access to credit, and rehabilitation of irrigation systems.

definition of the treatment variable. Moreover, I use the cohorts not exposed to the Green Revolution to conduct a placebo test and I show that the probability of having an arranged marriage did not change for them, providing additional support to my identification strategy.

I then use the Green Revolution to test the prediction that divorce among arranged marriage couples should increase as their insurance advantage decrease. First, I show that Indonesia and Turkey have had an increasing divorce rate among couples in arranged marriages, whereas the divorce rate among couples in love marriages has declined. Then, I study divorce in arranged marriages in the three first cohorts of my Indonesian sample. These individuals should have married for insurance but the Green Revolution lowered the insurance value of the arrangement. Indeed, I show that the Green Revolution changed the likelihood of divorcing. The cohort born before 1933 and exposed to the Green Revolution had a 18% lower probability of divorcing; however, this probability increased to 25% for the cohort born between 1933 and 1942, and to 29% for the cohort born between 1943 and 1952. The Green Revolution did not change significantly the probability of divorce for any cohort of the love marriage sample.

Overall, the empirical evidence is consistent with the hypothesis that arranged marriages disappear as the demand for informal insurance declines, specifically when countries experience economic transformations that raise the returns to education and lower the variance of income. The theoretical framework, however, delivers a richer set of predictions discussed in sections 4.1 to 4.4 (that are not alternative explanations to the evidence presented in this paper). On the one hand, the value of the informal insurance provided by the arranged marriages might have fallen as formal and informal arrangements appeared as substitutes (for instance, welfare programs or temporary migration) or as the type (and exposure to) of risks shifted as countries moved away from agriculture (parents no longer have an advantage in choosing the best insurance partner since they no longer observe the history of shocks).[4] On the other hand, the cost of belonging to this insurance arrangement might have increased as migration and urbanization have reduced the pool of potential insurance partners, increased the barriers to information flows and limited the enforceability of the contracts. However there are two alternative (and potentially complementary) explanations. One alternative is that in the past parents had full control over children (through controlling all the resources of the economy), but they have lost it as new (possibly more profitable) occupations become available. A second possibility, suggested by research within sociology and anthropology, is that love marriages result from the process of westernization and expansion of mass media that often accompanies economic development. I cannot entirely rule out these alternatives, but I discussed them in section 6.

Finally, this paper opens a new interesting and extensive research agenda. First, the case of South Asia presents a puzzle. The countries of this region (India, Pakistan and Bangladesh) are experiencing a similar transition away from agriculture, and they have adopted some of the new technology available in other countries; for instance, the higher yield seeds in the agricultural sector exploited in this paper. However, 95% of all marriages are still arranged. I discuss possible explanations for this puzzle in the paper. Second, modeling general equilibrium effects on the marriage markets might yield additional insights; for instance, dowry and bride-price payments might be needed to clear the markets. Lastly, the results on consumption smoothing and divorce indicate that the transition toward love marriages might have important welfare consequences for parents and children, and more generally, for economic growth as geographic and social mobility constrains are relaxed.

[4]As economies move away from agriculture, income risk is no longer only associated with weather, and parents lose their information advantage.

The rest of the paper is organized as follows. Section 2 briefly discusses the related literature. Section 3 shows the trends and patterns of marriage transition. Section 4.1 presents a simple model of marriage choice as a game between parents and one child. Section 4.2 presents the extension of the model that allows for divorce. Section 4.3 presents the second extension, where households have two children and different gender composition. Section 5 presents the empirical results. Finally, section 6 discusses alternative explanations, and briefly examines the case of South Asia and other future work.

2 Literature Review

Since Becker (1973, 1974), there has been an increasing interest in studying marital decisions and, more specifically, in studying the economics of the family. Despite the large literature developed since the 1970s, and perhaps surprisingly, there are only a few papers within economics that have studied the role of arranged marriages, which is a widespread practice in several regions of the world. Rosenzweig and Stark (1989) focus on India and analyze the strategies that rural households use to smooth consumption. They find that migration and marriage arrangements (mostly arranged by parents) contribute to mitigate variation in consumption expenditure. They show that parents choose partners living in distant villages for their daughters and use a spatial diversification strategy by marrying them into different areas with low weather covariance across them. They hypothesize that these marriage arrangements allow households to enter into an informal insurance contract and show that the variance in consumption expenditure decreases with the number of married daughters and the distance to the households where they are married. Munshi and Rosenzweig (2009) further show that low spatial and marital mobility in rural India are likely due to the existence of informal risk-sharing networks. Households marry within the sub-caste and remain living in rural areas as a way to show commitment with their network and strengthen their ties. Positive economic shocks that increase permanent income and that increase inequality within the insurance group, raise the likelihood that individuals leave the insurance networks, migrate to urban centers and marry outside the insurance network. The effect they find, however, is relatively small, leading them to conclude that these insurance networks generate large benefits relative to the outside option. To my knowledge, relative to this literature, this paper is the first to show the transition away from arranged marriages outside South Asia, to propose an explanation for it and to provide empirical evidence consistent with a decline in the insurance motive.

Beyond the literature in economics, there is a large research in sociology and anthropology that asserts that families use the marriage of their children as a way to create alliances with other groups and to strengthen their social ties within their communities. Some examples from the literature on anthropology that has focused on Indonesia (where the main analysis of this paper is carried out) are the following:

> "Marriage, in adat law, is in varying degrees a matter of kinship group, community and personal concern. It is also a matter of social status. Marriage is the means by which the organized relationship groups which form autonomous communities maintain their existence. Social classes maintain themselves through well-regulated marriages, and hence the tie-up between marriage and social status.[...] Fellow members aid each other reciprocally. And groups, particularly kin groups, and exogamous sub-clans, are in a regular exchange of goods, which is linked to the

exchange of women." (Ter Harr, 1948).[5]

"Parental marriage arrangement in Java must be seen not in terms of kinship organization as such, but as an aspect of the economic and prestige systems of the larger society, and as a function of the internal authority structure of the elementary family. For the choice of spouse, serves the interests of the parents primarily, by expanding the range of their social ties, or consolidating those already existing, and by validating their social rank in their community." (Geertz, 1961).

"Adat never protects individual interests but guarantees in first place interests appertaining to the group. The settlement of a marriage should be regarded as an agreement between two families. Marriage and issue do not exist to further the happiness of the individual; they have a very different meaning: they are institutions which help to maintain the existence of the clan." (Vreede-de Stuers, 1960).

Similar research can be found for other countries, for instance:

"Leaving aside divorce-dissolved families, couples who marry after free courtship are less likely permanently to be obliged to provide material and emotional care for their relatives and in-laws, in particular the husband's parents." (Korea: Chang, 1997).

"At the micro level, especially in developing countries like Pakistan, the family remains centrally responsible for providing food and sustenance, offering also protection and safety to individuals, particularly in childhood and old age. The family as a supreme institution, however, then also dominates individual agency and asserts its will over choices in marital selection, thus potentially undermining individual emotions and causing hardship. Families promote such marriages where they perceive the possibility of gaining certain types of benefits or various forms of security." (Pakistan: Zaman, 2008).

"The fact that lineages and clans are widespread does not differentiate Africa, except in degree, from many other culture areas. They are common, however, and thus the choice of spouse, gift exchanges at marriage, and the subsequent attention paid to marital behavior by the clan has a corporate character." (Sub-Saharan Africa: Goode, 1970).

"The maintenance of the caste system [...] depended completely upon the arranged marriage. Maintenance of the caste was too important a matter to be left to the young. [...] in India it developed not only among the wealthy, who could afford early marriages and whose union might mark an alliance between two families, but also among the poor, who had nothing to share but their debts." (India: Goode, 1970).

In contrast with the little attention paid to arranged marriages, there is an important and large strand of the literature of development economics that has focused on understanding the (ex-ante and ex-post) mechanisms used to cope with risk, especially in rural areas. This literature has found overwhelming evidence of large needs of insurance, which often lead to inefficient choices in other spheres, creating poverty-traps and exacerbating inequalities (Banerjee and Duflo, 2007; Morduch, 1994, 1995). Among the ex-ante strategies we can find shared tenancy or share-cropping, which has been shown to create inefficient investments but might allow risk-sharing between a landlord and a tenant (Stiglitz, 1974; Ackerberg and Botticini, 2002); asset and occupational diversification, and investment in less risky (less profitable) assets (Rosenzweig and Binswanger, 1993); savings (Paxson, 1992; Calomaris and Rajaraman, 1998; Klonner, 2003); formal insurance, e.g. weather based insurance, although there is large evidence of low take-up rates when available (Hazell et al., 2010; Cole et al., 2013; Giné, Townsend and Vickery, 2008); and location diversification through temporary or permanent migration (Lucas and Stark, 1985; Rosenzweig and Stark, 1989; de la Briere et al., 2002; Morten, 2013). Once shocks are realized, individuals also engage in activities to mitigate their

[5] Adat law is the term used to refer to customary law.

effects on consumption. Ex-post mechanisms include borrowing (Udry, 1994), selling assets or dis-saving (Rosenzweig and Wolpin, 1993; Fafchamps, Udry and Czukas, 1998), transfers (Townsend, 1994; Ravallion and Chaudhuri, 1997) and other labor supply adjustments (Jayachandran, 2006). Despite the efforts to mitigate risk, there is substantial evidence that these mechanisms are not sufficient.[6] Relative to this large literature, this paper provides some evidence on one risk sharing mechanism, arranged marriages, and how it has evolved and disappeared as economic growth takes place. However, I leave open the question of the long-run consequences for consumption smoothing of parents and children.

This literature has also tried to quantify the extend of the insurance in both developed and developing countries (how effective are the ex-post informal insurance mechanisms (transfers) at insuring consumption). Mace (1991) and Cochrane (1991) test the hypothesis of full insurance using data from the USA. Both papers reject the hypothesis of efficient risk sharing under homogeneous preferences. Townsend (1994) and Ravallion and Chaudhuri (1997) perform a similar test in rural India using data from ICRISAT and assuming that the village is the relevant unit for risk sharing. Although both papers find co-movement between individual consumption and aggregate consumption, they reject the full insurance hypothesis since they find that income does matter for consumption decisions. Mazzocco and Saini (2012) relax the assumption of homogenous risk preferences allowing for heterogeneity in the taste for risk and perform a test of full insurance at the sub-caste level. They are unable to reject efficient risk sharing at the sub-caste level under heterogeneous preferences. Borrowing from this literature, I implement the standard test of full insurance to show that the consumption of arranged marriage does not vary with their income, whereas the consumption of love marriages does. However, my results also suggest that the village or town might not be the appropriate insurance unit.

This study also contributes directly to the larger literature initiated by Gary Becker (1973, 1974, 1991) that studies the determinants of marriage. As summarized by Browning, Chiappori and Weiss (2014), marriage exists due to potential gains shared by the couple: sharing of a public good, division of labor to exploit comparative advantage, extension of credit and coordination of investment activities, and risk pooling, among others. Within this large literature, this paper is close to Hess (2004), who studies the decision to marry and divorce in the presence of incomplete markets. He constructs and tests a model of individual marriage decision where men and women insure each other, and he shows that joint economic characteristics from the beginning of the marriage have a sizable impact on marital duration. Shore (2010) presents evidence supporting the risk-sharing component of marriage. He shows that while individuals face more idiosyncratic risk in bad times, households do not. In contrast to their work, this paper focuses on the insurance gains of both spouses *and* their families (parents and parents-in-law) and studies the transition from marriage as a decision of the family to marriage as a decision of the couple. Within the literature of development economics there is has been an effort to understand the determinants of marriage in developing countries, particularly in rural areas (Fafchamps and Quisumbing, 2008). Among the main factors behind the formation of households, they emphasize the role of insurance, savings and capital accumulation, and parental involvement. Relative to this literature, this paper presents supportive evidence on the insurance gains from arranged marriages and further shows that economic growth has eroded the value for these informal insurance arrangements.

[6] Rose (1999) finds that drought increases mortality among Indian girls; Jacoby and Skoufias (1997) find that child labor, and thereby school attendance, play a significant role in the self-insurance strategy of rural household in India; Gertler and Gruber (2002) find that while families are able to full insure minor illness, they are not able to insure illnesses that limit their ability to physically perform activities of daily living; Maccini and Yang (2009) show that women who experienced drought as young children are shorter, poorer, and obtain less education. These are only a few examples of the short-run and long-run consequences of imperfect insurance markets and imperfect risk coping mechanisms.

Finally, although I leave for future work the question of long run consequences of the transition to love marriages, Edlund and Lagerlof (2006) propose a model where love marriages might cause economic growth.

3 Stylized Facts

The literature on sociology, anthropology and psychology (Goode, 1970; Buunk et al., 2008; Jones, 2010; among others) suggests that arranged marriages originated as a strategy of families to form alliances with other families, groups or clans.[7] Arranged marriages have existed in most societies throughout history. In Europe, arranged marriages disappeared as the Catholic Church consolidated its place as the main religion. Anthropologist Jack Goody (1983) documents that arranged marriages were common among the ancient Greeks, Romans and Anglosaxon tribes until the rise of the Catholic Church, which favored self-choice marriage and monogamy. The goal of these rules was to limit the ability of families to form alliances (and, thus, to limit their ability to increase wealth) and to limit the number of legitimate heirs in order to divert inheritance toward the Church. This evolution was not monotonic nor uniform; however, by the late medieval period, the nuclear monogamous and self-chosen marriage was dominant in Europe (Greif 2006), except among the wealthiest class, which continued to arrange marriages for their children until the dawn of the Industrial Revolution (Goode, 1970).

In Asia and Africa, arranged marriages continued to be the dominant marriage institution until recent decades. In a companion paper (Rubio, 2013), I provide extensive and detailed analysis of the transition for eighteen countries: Turkey, Saudi Arabia, Israel, Japan, Korea, China, Taiwan, Indonesia, Malaysia, Cambodia, Vietnam, Sri Lanka, Nepal, Togo, Ghana, India, Pakistan and Bangladesh. Figure 1 shows these trends by region (Middle East and Africa, East Asia, South East Asia and South Asia). For the first three regions, we observe a clear trend toward the disappearance of arranged marriages, although at different rates. The exception is South Asia (India, Bangladesh and Pakistan), where arranged marriages are still the most common form of marriage. However, a closer examination of India suggests that urban areas have started the transition to love marriages, increasing from 5% to 10%.

The next four figures show some of the main correlates of the decline of arranged marriages. Figure 2 shows that the decline in the share of arranged marriages across cohorts is correlated with the increase in educational attainment. Furthermore, education and arranged marriages are negatively correlated even within cohorts and countries, ruling out the possibility that the time series correlation is exclusively driven by a common unobserved time trend (Figure 3). Figure 4 shows that the decline in arranged marriages is correlated with the increase in the percentage of women working in the formal labor market, and figures 5 and 6 indicate that the decline in agriculture and the increase in urbanization are also associated with lower shares of arranged marriages. Although not shown here, most of these results also hold when using a regression analysis that includes all the variables.[8] Overall, the correlations suggest that economic growth has been associated with a dramatic change in the formation of households.

[7] Another example found in the evolutionary psychology literature also supports the hypothesis posed by sociologists and anthropologists: "Parents may have a relatively stronger preference for children's mates with characteristics suggesting high parental investment and cooperation with the in-group, whereas children may have a relatively stronger preference for mates with characteristics signaling heritable fitness." (Buunk et al., 2008)

[8] Results available upon request.

4 The Model

The fundamental assumption of the model is that arranged marriages are used as an insurance mechanism. The model emphasizes a potential trade-off between insurance gains from arranged marriages and returns outside the social network (the outside option).[9] The outside option is represented by unconstrained returns to education; if individuals give up the arranged marriage, they have the possibility of moving geographically (and thus accessing a broader set of occupations) and/or finding a partner with higher education/income.[10] I start by analyzing a two-period model for a household with only one child which captures this trade-off.[11] I then extend the the model to allow for the possibility of divorce. Finally, I modify model to analyze how the number of children and their gender composition might affect parental choices.

4.1 One-child Model: Setup

There are two periods and each household has two agents: parents (who act as one agent) and one child, denoted by the subscript f and k, respectively. Each agent maximizes a quadratic utility function, $u(c_{t,i}) = c_{t,i} - \frac{d_i}{2}c_{t,i}^2$, $i = f, k$, $t = 1, 2$, where $c_{i,t}$ is consumption of agent i at time t, d_i is the parameter that captures her degree of risk aversion and it is bounded such that $u(c_{i,t}) > 0$, $u'(c_{i,t}) > 0$ and $u''(c_{i,t}) < 0$ in the relevant region in which $c_{i,t}$ takes values. Parents and child have an income endowment normalized to 1 in each period.

In period 1, parents choose investment in education λ_k for the child and the level of effort, $e \in \{0, 1\}$, they exert to find her a partner. The first period budget constraint for parents is given by $c_f = 1 - p\lambda_k - e_{high}I(e = 1)$, where p is the price of education, $I(e = 1)$ is an indicator variable taking the value of one if parents choose high effort and e_{high} is its cost. In this period, the child consumes what she produces. In period 2, the child receives $x_k\lambda_k + \delta_k$, where x_k are the known returns to her education λ_k and faces a shock $\delta_k \sim N(0, \sigma_\delta^2)$. Formally, in this model the average returns to schooling are different by type of marriage, $x_{k,h}$, $h = L, A$. The shock, however, is the same regardless of the level of education or type of marriage. Parents receive a share $0 < \varphi < 1$ from the returns to schooling of the child (this assumes commitment between parents and children under both type of marriages; the data supports this assumption, table 1 at the bottom shows evidence on transfers from and to parents and parents-in-law for both types of marriages).

All children marry at the beginning of the second period, and within marriage the child shares resources with her spouse equally. Under these assumptions, the consumption in the second period for each agent is given by:

[9]The recent literature in development economics has modeled this problem in the context of a limited commitment model (Coate and Ravallion, 1993; Lingon, Thomas and Worral, 2002), where households participate in the agreement as long as they receive at least a reservation value (equivalent to the outside option being lower than the net gain from the risk-sharing agreement). The model proposed in this section omits the potential limited commitment problem; however, a similar intuition should follow if we relax this assumption.

[10]To solve the model, I assume assortative matching in terms of education level which might be interpreted as children finding a partner with higher education/income.

[11]The model also captures other economic changes experienced throughout this period and the numerical simulations try to show their relative importance: (i) Changes in the risk profile (as countries move from agriculture to manufacture and services, and as welfare programs are introduced by governments) can be studied by changing the variance and covariance (across agents) of the shock to returns to schooling; (ii) The increasing cost of informal insurance (as migration and urbanization reduce the pool of potential insurance partners, increase the barriers to information flows and limit the enforcement of the agreements) is introduced through an effort cost that parents exert for finding a partner for their child; (iii) The outside option is captured by assuming potential higher returns to education in love marriages (relative to arranged marriages); and (iv) other changes in the marriage markets are introduced through a match or love term drawn from a distribution that differs by type of marriage.

$$c_{k,h} = 1 + (1-\varphi)(\frac{x_{k,h}\lambda_{k,h} + x_{s,h}\lambda_{s,h}}{2} + \frac{\delta_k + \delta_s}{2}) \quad (1)$$

$$c_{f,h} = 1 + \varphi(\frac{x_{k,h}\lambda_{k,h} + x_{s,h}\lambda_{s,h}}{2} + \frac{\delta_k + \delta_s}{2}) \quad (2)$$

where $x_{s,h}\lambda_{s,h}+\delta_s$ is the income of the child's spouse and assortative matching is assumed between spouses in terms of education.[12] The child receives additional utility from a love term, $u(c_{k,h}) + \alpha_h$, $h = A, L$, which comes from a known distribution: the cdf in the arranged marriage market is denoted by $\alpha_A \sim F_A(\alpha)$, and the cdf in the love marriage market is represented by $\alpha_L \sim F_L(\alpha)$.

The effort of parents in the first period determines the insurance quality of the partner in the second period. I define insurance quality as the correlation between the child's shock and her spouse's shock, $\varrho_{ks}\left(I(e=1), I(L=1)\right)$, where $I(L=1)$ is an indicator variable taking the value of one if the child chose the love marriage: (i) If parents exert high effort, $e = 1$, and the child accepts the arranged marriage, $L = 0$, she and her spouse have a perfectly negatively correlated income; (ii) If parents exert low effort, $e = 0$, and the child accepts the arranged marriage, $L = 0$, the negative correlation between spouses' shocks is less than perfect — for simplicity I assume that it is $\varrho_{ks}\left(I(e=0), I(L=0)\right) = 0$; (iii) If the child decides to find her own mate in the love marriage market, $L = 1$, the correlation with her spouse's income might be positive, negative or zero, regardless of the effort of the parents — for solving of the model I set it equal to zero.

4.2 Analysis and Discussion

The model is solved backwards. Starting in period 2, parents and children calculate their expected utility for a given level of education, a given effort and an expected love term. Anticipating the decision of the child in period 2, parents choose effort and education in period 1 by solving the following maximization problem:

$$\underset{\lambda_{k,h}, e\in\{0,1\}}{Max} \; u(c_f) + \beta E[u(c_f)] \quad (3)$$

$$s.t. \quad c_f = 1 - p\lambda_{k,h} - e_{high}I(e=1)$$

where:

$$\begin{aligned} E[u(c_{f,h})] \;=\; & \left[1 + \varphi\left(\frac{x_{k,h}\lambda^*_{k,h} + x_{s,h}\lambda_{s,h}}{2}\right)\right] - \frac{d}{2}\left[1 + \varphi\left(\frac{x_{k,h}\lambda^*_{k,h} + x_{s,h}\lambda_{s,h}}{2}\right)\right]^2 \\ & - \frac{d}{2}\varphi^2\sigma^2_\delta\left(\frac{1 + \varrho_{ks}\left(I(e=1), I(L=1)\right)}{2}\right) \end{aligned} \quad (4)$$

The first order condition delivers:[13]

$$\lambda(e)^*_{k,h} = \frac{(\beta\varphi x_{k,h} - 2p)(1-d) - 2pde_{high}I(e=1)}{d(2p^2 + \beta\varphi^2 x^2_{k,h})} \quad (5)$$

Parents invest in the child's education if the discounted share of returns they receive ($\beta\varphi x_{k,h}$) is higher

[12] This assumption allows me to find the optimal education level for the child in terms of the parameters of the model; otherwise, the optimal education for the child will depend on the expected education (and return) of her spouse. This assumption might be relaxed to analyze other cases.

[13] In equilibrium $x_{k,h} = x_{s,h}$, by the assumption of assortative matching. Therefore the FOC can be simplified and expressed only in terms of $x_{k,h}$.

than the foregone consumption in the first period $(2p)$.[14] Parents also face a trade-off $(2pde_{high}I(e=1))$ between investing in education and looking for a high quality insurance partner for their child.[15]

The optimal effort is chosen based on the comparison of the expected utility under each scenario; high effort, $e = 1$, is optimal if its present discounted value of consumption is higher than choosing the alternative (suppressing the subscripts k and h):

$$[1 - p\lambda^*(e=1) - e_{high}] - \frac{d}{2}\left[1 - p\lambda^*(e=1) - e_{high}\right]^2 + \beta\left\{[1 + \varphi x\lambda^*(e=1)] - \frac{d}{2}\left[1 + \varphi x\lambda^*(e=1)\right]^2\right\} > \quad (6)$$
$$[1 - p\lambda^*(e=0)] - \frac{d}{2}\left[1 - p\lambda^*(e=0)\right]^2 + \beta\left\{[1 + \varphi x\lambda^*(e=0)] - \frac{d}{2}\left[1 + \varphi x\lambda^*(e=0)\right]^2 - \frac{d}{2}\varphi^2\sigma_{i\delta}^2\right\}$$

High effort decreases consumption in the first period and decreases education $(\frac{\partial\lambda}{\partial e_{high}}\mid_{I(e=1)} < 0)$ (and therefore consumption in the second period), but it is optimal as long as the child accepts the arranged marriage and the gains in utility from offsetting the income shock $(\frac{d}{2}\varphi^2\sigma_\delta^2)$ are large enough to compensate for the loss in consumption in both periods.

In this model all children marry; they decide at the beginning of period 2 either to accept the arranged marriage or not based on the comparison of the expected utility under each marriage (suppressing the subscript k):

$$E[u(c_{k,L}) + \alpha_L] - E[u(c_{k,A}) + \alpha_A] = (1-\varphi)\left(x_L\lambda_L^* - x_A\lambda_A^*\right)\left[(1-d_k) - \frac{d_k}{2}(1-\varphi)\right. \quad (7)$$
$$\left.(x_L\lambda_L^* + x_A\lambda_A^*)\right] - d_k\frac{(1-\varphi)^2}{4}\sigma_\delta^2\left[\varrho_{ks}\left(e, I(L=1)\right) - \varrho_{ks}\left(e, I(L=0)\right]\right. + E\left(\alpha_L\right) - E\left(\alpha_A\right) > 0$$

Using these results and the assumptions outlined, we can summarize the main implications of the model in the following proposition.

Proposition 1. *Parents and children receive benefits from insurance (and thus from arranged marriages) and from the returns to education. The child, however, receives additional utility from the love term; she might be willing to give up insurance in order to find a love mate. In contrast, parents do not receive utility from this love term, generating a wedge between the child and the parents. Ceteris paribus, love marriage is preferred when (among others):*

(i) $(x_L - x_A) > 0$, *the returns to education are higher or increasing in love marriages. For a given level of education, higher* ***unconstrained*** *returns increase the probability that the child chooses the love marriage. In turn, parents internalize it, decrease effort, which produces two additional effects. It further increases education, which increases the value of the outside option. It also decreases the insurance quality of the arranged marriage mate, decreasing the insurance benefits of the arrangements.*

(ii) *And,* σ_δ^2, *the size of the shock decreases. For a given level of education, a decrease in the size of*

[14]The optimal education level is increasing in the returns to education and on the share that parents receive as long as $\frac{2p}{\beta}\frac{[1+d(e_{high}-1)]}{(1-d)} < \varphi x_{k,h} < \frac{4p}{\beta}\frac{[1+d(e_{high}-1)]}{(1-d)}$. For the rest of the analysis, I assume that the returns to education fall within this range in order to derive comparative statics.

[15]The introduction of the effort cost mechanically introduces a trade-off between investing in education and finding an insurance partner for the child. This trade-off might be assumed away by setting $e_{high} = 0$, the main results still follow.

the shock decreases the probability that the parents will exert high effort; in turn, this lowers the insurance quality of the arranged marriage partner. It also increases the investment in education, λ_k, effectively increasing the outside option of the child. For the child, the insurance advantage of the arranged marriage also disappears.

Proof. *The appendix shows the analytical proofs for changes in $(x_L - x_A)$ and σ_δ^2. In addition it shows analytical results for changes in $\varrho_{ks}(e, I(L=1)) - \varrho_{ks}(e, I(L=0)$, d_k, d, e_{high}, and $E(\alpha_L) - E(\alpha_A)$.*

Figure 7 illustrates the results described in proposition 1 and some additional comparative statics (discussed in appendix 8.1). The graphs are generated by simulating the decisions of the two players (parents and child) for a given set of parameters; the simulations are performed in Matlab using the solution to the model (optimal educational investment, optimal effort and optimal marriage choice).[16] Each graph plots the probability of accepting the arranged marriage (of the child) and the probability of exerting high effort (of the parents) against each of the relevant parameters. Row 1, column 1, illustrates how the probability of love marriage increases as $x_L - x_A$ increases. Intuitively, the increase of this term represents the rise of the outside option, and it decreases the probability that parents exert high effort (lowering the value of the informal insurance) and the probability that the child accepts the arranged marriage (for a given quality of arranged marriage partner and a given education). The next column shows that as $\uparrow \sigma_{i\delta}^2$ (the expected size of the shock increases), the parents' probability of exerting high effort increases, lowering the education of the child (and her outside option), increasing the insurance quality of the partner and thus increasing the probability that the child chooses the arranged marriage. For the child, holding everything else constant, the benefit of insurance dominates the outside option, and she will likely choose the arranged marriage.

The rest of the graphs present additional predictions which are not tested in the data. The third column illustrates that as $\uparrow \varrho_{ks}(I(e=1), I(L=1))$ (income correlation between spouses increases) in the arranged marriage market, the insurance gains vanish and both probabilities decrease. In contrast, for a given insurance benefit in the arranged marriages, more risk-averse children ($\uparrow d_k$, row 1, column 3) are more likely to accept the arranged marriage and, in turn, parents respond by increasing the high effort search (since it delivers the best insurance partner). Finally, as the expected love term in love marriage increases relative to the expected love term in arranged marriages ($\uparrow E(\alpha_L) - E(\alpha_A)$, row 2, column 2), the speed of the transition increases. From the side of the parents, figure 7 shows the response to changes in the parameters that affect their decisions. Row 2, column 3 shows that more risk-averse parents ($\uparrow d$) endogenously choose lower levels of education and exert high effort, decreasing the value of the outside option of the child, and inducing her to accept the arranged marriage. The price of education, p, has a similar impact; row 3, column 1 shows that when education is very costly (for example, there are no schools available), parents invest in low levels of education and children are better off accepting the arranged marriage. The cost of high effort can be interpreted as a measure that captures changes in the social network (search cost). The increasing cost of high effort, $\uparrow e_{high}$, affects the parents' consumption in the first period and decreases the education of the children, affecting the transfer that parents receive in period two. As the cost increases (row 3, column 2), parents prefer to exert low effort, increasing the education investment, and children switch faster to the love

[16]The parameters used for the simulations are $p \in [0.1, 0.3]$, $\beta \in [0.8, 1]$, $\varphi \in [0.05, 0.3]$, $d \in [0.15, 0.65]$, $d_k \in [0.01, 0.55]$, $e_{high} \in [0.01, 0.25]$, $e_{low} = 0$, $\sigma_\delta^2 \in [0.01, 1.2]$, *correlation* $\in [-1, 1]$, $\alpha_L \in [0.01, 0.2]$, and $\alpha_A \in [0.001, 0.1]$; x_L and x_A are bounded such that education is increasing in the returns.

marriage market.

The analysis of section 5 will only test the predictions stated in proposition 1.

4.3 Extending the Model to Include Divorce

The goal of this section is to understand how divorce behavior differs by type of marriage as the insurance advantage of arranged marriages vanishes. This extension is motivated by large divorce rate in some Southeast Asian countries. In order to derive the intuition, I allow for a third period when the divorce decision takes place.

The sequence of decisions remains unchanged for the first two periods. At the beginning of period 3, the child observes the realization of the love term, α, and decides whether to remain married or not. If she divorces, she will face the realization of her shock and pay a utility cost $\phi > 0$ in period 3; and in the next period, she will find a new partner in the love market regardless of the previous type of marriage. Therefore, the child will divorce if:

$$\alpha_h + u\left(c_k\right)^{M,h} + \beta\left\{E\left[u\left(c_k\right)\right]^{M,h} + \alpha_h\right\} < u\left(c_k\right)^{D} - \phi + \beta\left\{E\left[u\left(c_k\right)\right]^{M,L} + E\left(\alpha_L\right)\right\},\ h = L, A \tag{8}$$

M represents the utility of married individuals, D represents the utility of divorced individuals, h refers to the type of marriage, L refers to love marriage and A to arranged marriage; α_h is the realized love term for the married individual in type of marriage $h = L$, A, and α_L is the love term drawn from $F_L(\alpha)$. This expression is simplified in the appendix and used to derive the thresholds for divorce by type of marriage.

Since the distribution of α_h differs by marriage, the probability of divorce depends directly on the distribution of the love term in each marriage market. Let us define the expressions (see appendix 8.2 for details on these thresholds):

(i) $$\hat{\alpha}_L = (1+\beta)^{-1}\left[-\phi + \tfrac{d_k}{2}\left(1-\varphi\right)^2\sigma_\delta^2\left[\tfrac{\varrho_{ks}(e,\, I(L=1))}{2} - \tfrac{1}{2}\right] + \beta E\left(\alpha_L\right)\right]$$

(ii) $$\hat{\alpha}_A = (1+\beta)^{-1}\left[-\phi - \tfrac{d_k\beta}{2}\tfrac{(1-\varphi)^2}{4}\sigma_\delta^2\left[\varrho_{ks}\left(e,\ I(L=1)\right) - \varrho_{ks}\left(e,\ I(L=0)\right] + \tfrac{d_k}{2}\left(1-\varphi\right)^2\sigma_\delta^2\left[\tfrac{\varrho_{ks}(e,\, I(L=0))}{2} - \tfrac{1}{2}\right] + \beta E\left(\alpha_A\right)\right]$$

Recall that $\alpha_L \sim F_L(\alpha)$ and $\alpha_A \sim F_A(\alpha)$, delivering the following probabilities of divorce (divorce occurs for any draw, α_h, that falls below the thresholds defined above):

$$P^{D,L} = \int_{-\infty}^{\hat{\alpha}_L} dF_L(\alpha)d\alpha \ \ and \ \ P^{D,A} = \int_{-\infty}^{\hat{\alpha}_A} dF_A(\alpha)d\alpha \tag{9}$$

The solution to the model is found in a similar way as before. Once the probabilities of divorce are calculated for each type of marriage, the child uses them to calculate the expected utility for each type of marriage in period 2, and the parents incorporate them into the optimal choices (education and effort) of period 1. The appendix 8.2 shows the expressions determining these choices.

Proposition 2. *For arranged marriages, ceteris paribus, divorce increases as the gains from insurance* *disappear:* $\downarrow [\varrho_{ks}(e,\, I(L=1)) - \varrho_{ks}(e,\, I(L=0)]$.

Proof. *The threshold,* $\hat{\alpha}_A$*, increases as* $\downarrow [\varrho_{ks}(e,\, I(L=1)) - \varrho_{ks}(e,\, I(L=0)]$*. This leads to an increase in* $P^{D,A}$ *from equation 9.*

If arranged marriages provide more insurance, $\varrho_{ks}(e,\, I(L=1)) > \varrho_{ks}(e,\, I(L=0)$, their divorce threshold will be smaller $\hat{\alpha}_A < \hat{\alpha}_L$. In countries where the cost of divorce (ϕ) is the same by type of marriage, we should expect $P^{D,A} > P^{D,L}$ only if $F_L(\alpha)$ first order stochastically dominates (FSD) $F_A(\alpha)$ and the difference between thresholds, $\hat{\alpha}_A - \hat{\alpha}_L$, is small. If the threshold difference is large, or if $F_A(\alpha)$ FSD $F_L(\alpha)$, then we should expect the opposite result. Notice, however, that the risk aversion of the child, d_k, will determine how important the variances and covariance of the shock are. Less risk-averse children will place a higher weight on the love term. β plays a similar role; impatient children will put a higher weight on the dis-utility generated by being single during period 3. While these features are interesting, section 5 will only test the prediction of proposition 2.

4.4 Extending the Model to Two Children

I extend the model to consider the case in which parents have two children. This extension enables me to examine two important dimensions that are assumed away in the base model: the role of the size of the network and the effect of the gender composition of the children (within the household, abstracting from general equilibrium effects on the marriage markets). As the number of children increases, the size and quality of the social network play a crucial role; parents must take into account that the households where their children may be married might have correlated shocks. If the social network is small (in this context, equivalent to having insurance partners with positively correlated shocks), parents have incentives to arrange marriages for only some children.

For the rest of the section I assume that all households have two children, differing only in the gender composition: (i) 2 boys; (ii) 1 girl and 1 boy; (iii) or 2 girls. In period 1, parents invest in the education of both children and decide the amount of effort exerted looking for a partner for each. The first period budget constraint is now given by:

$$c_f = 1 - gp_g\left(\frac{1}{j}\sum_j \lambda_{g,j}\right) - bp_b\left(\frac{1}{n}\sum_n \lambda_{b,n}\right) - g(\frac{1}{j}\sum_j e_j^g I(e_{j,g}=1)) - b(\frac{1}{n}\sum_n e_n^b I(e_{n,b}=1)) \qquad (10)$$

where $j = 0, 1, 2$ and $n = 0, 1, 2$ are the number of girls and boys, respectively; $N_k = j + n$ is the total number of children; $g = \frac{j}{N_k}$ is the share of girls and $b = \frac{n}{N_k}$ is the share of boys; $e_{g,,j} \in \{0,1\}$ is the effort for girl j with cost e_j^g and $e_{b,n} \in \{0,1\}$ is the effort for boy n with cost e_n^b; $I(e_{j,g}=1)$ is an indicator variable that takes the value of 1 if parents choose to exert high effort for girl j; and $I(e_{n,b}=1)$ has a similar interpretation for boy n. Children are homogeneous within gender but heterogeneous between gender in the price of education ($p_g \neq p_b$) and the returns to schooling ($x_g \neq x_b$).[17] Again, the returns to schooling differ

[17] We might interpret the differences in prices as boys and girls having different opportunity cost of studying; for example, girls might have a lower opportunity cost of being taking away from home or from agricultural production (in the societies where female labor is less used for agricultural production). The differences in returns to schooling might be considered in a similar way; in agricultural societies, boys might have an advantage due to larger returns to physical strength, so as countries move away from agriculture the differences in returns to education might be reduced. The main goal of this section is to explore

by type of marriage, $h = A, L$; for notational simplicity, I am omitting the subscript h.

Under these assumptions, the choice of education for the children depends on the gender composition of the family. I focus here on the case of one boy and one girl; the other two cases can be analyzed in a similar manner. The problem faced by the parents in the first period can be re-written as:

$$\underset{\lambda_{g,j},\lambda_{b,n},e_j\in\{0,1\},e_n\in\{0,1\}}{Max} u(c_f) + \beta E[u(c_f)] \tag{11}$$
$$s.t. \quad c_f = 1 - \tfrac{1}{2}p_g\lambda_{g,1} - \tfrac{1}{2}p_b\lambda_{b,1} - \tfrac{1}{2}e_1^g I(e_{g,1}=1) - \tfrac{1}{2}e_1^b I(e_{b,1}=1)$$

For given effort levels $e_{g,1}$ and $e_{b,1}$, the first order conditions for $\lambda_{g,1}$ and $\lambda_{b,1}$ determine the parents' optimal investment in education for boys and girls:

$$if\ \frac{x_g}{p_g} > \frac{x_b}{p_b} \Rightarrow \lambda_{g,1}^* = \frac{(1-d)\left(\beta\varphi x_g - 2p_g\right) - 2dp_g e}{dg\left(2p_g^2 + \beta\varphi^2 x_g^2\right)},\ \lambda_{b,1}^* = 0 \tag{12}$$

$$if\ \frac{x_g}{p_g} < \frac{x_b}{p_b} \Rightarrow \lambda_{b,1}^* = \frac{(1-d)\left(\beta\varphi x_b - 2p_b\right) - 2dp_b e}{db\left(2p_b^2 + \beta\varphi^2 x_b^2\right)},\ \lambda_{g,1}^* = 0 \tag{13}$$

where $e = \frac{1}{2}e_1^g I(e_{g,1}=1) + \frac{1}{2}e_1^b I(e_{b,1}=1)$.

In the case of households with two boys or two girls, we might expect a priori that parents provide the same level of education to both children (they are homogeneous within gender); however, the final choice of education might be asymmetric and it will depend on the marriage choice of each child (see discussion in appendix 8.3).

The model is solved backwards starting in the second period. Parents and children calculate their expected utility for a given level of education and effort. The expected utility of parents is given by (let $e_{g,1}$ and $e_{b,1}$ denote the chosen effort):

$$\begin{aligned} E[u(c_f)] \;=\; & \left\{ \left[1 + \varphi\left(gx_g\lambda_g^* + bx_b\lambda_b^*\right)\right] - \tfrac{d}{2}\left[1 + \varphi\left(gx_g\lambda_g^* + bx_b\lambda_b^*\right)\right]^2 \right. \\ & -\frac{d}{2}\sigma_\delta^2 \left\{ \varphi^2\left(\frac{1 + \varrho_{g_1 s}\left(I(e_{g,1}=1),\, I(L=1)\right)}{2}\right) + \varphi^2\left(\frac{1 + \varrho_{b_1 s}\left(I(e_{b,1}=1),\, I(L=1)\right)}{2}\right) \right. \\ & \left. +\varphi^2 \varrho_{g_1,b_1}(e_{g,j}, e_{b,i}) \right\} \end{aligned} \tag{14}$$

where $\varrho_{g_1,b_1}(e_{g,j}, e_{b,i}) = \varrho_{g_1 b_1} + \varrho_{g_1 s_2} + \varrho_{b_1 s_1} + \varrho_{s_1 s_2}$.[18] This term captures the correlation between the households where the children are married, and it depends on: (i) the effort level exerted for each child (recall that effort determines the insurance quality of the partner proposed by the parents); and (ii) the type of marriage chosen by each child. In contrast to parents, each child still decides based on 7, which does not depend on $\varrho_{g_1,b_1}(e_{g,j}, e_{b,i})$.

In the extreme case in which the parents belong to a very small network, i.e., they have access to only one potential insurance partner (only one available household), arranging the marriage of both children into this

gender differences that in the absence of differences in cost and/or returns are not present, leading to a less interesting case for analysis. In the absence of gender differences, the analysis from the previous section can be directly applied provided that the budget constraint is properly adjusted to account for more costly children (as there are more children within a household).

[18] $\varrho_{g_1 b_1}$ is the income correlation between the two children; $\varrho_{g_1 s_2}$ is the income correlation between the first child and the spouse of the second child; $\varrho_{b_1 s_1}$ is the income correlation between the second child and the spouse of the first child; and $\varrho_{s_1 s_2}$ is the income correlation between the spouses of the two children.

household will increase the dis-utility term (the last term of 14, $\varrho_{g_1,b_1}(e_{g,j}, e_{b,i})$) instead of providing more insurance;[19] furthermore, as the number of children increases, the concern of a small social network increases as well (see appendix 8.3 for proof).

Proposition 3. *If* $\varrho_{g_j,b_n}(e_{g,j} = 1, e_{b,n} = 1) \Rightarrow \varrho_{k,s} = 1 \; k = g_j, b_n, \; s = s_j, s_n$ *(the most constrained case, only one potential insurance partner, each component of* $\varrho_{g_1,b_1}(e_{g,j}, e_{b,i})$ *has a positive correlation equal to 1) and* $\varrho_{g_j,b_i}(e_{g,j} = 0, e_{b,i} = 1) = \varrho_{g_j,b_i}(e_{g,j} = 0, e_{b,i} = 0) = 0$*, then, ceteris paribus, parents exert high effort for (offer the arranged marriage) and give no education to the child with the lowest net return in the labor market. Parents invest in positive education for the child with the highest net return in the labor market and exert low effort for her (the appendix 8.3 and figure* 13 *show a more detailed analysis for different gender composition).*

Proof. *Appendix.*

In summary, the results suggest that parents use education investment and effort to induce children to accept the arranged marriage. By reducing education (for the child with the lowest net returns), parents are effectively reducing her outside option. When the child with a low outside option is offered the possibility of entering into the insurance arrangement by marrying a high insurance quality partner (through parents exerting high effort for finding her a partner), she will likely accept it. These results depend crucially on the assumption of limited or small social network, which imply that the households where the children would be married (under arranged marriages) have high income correlation.

Corollary 3.1. *As* $\varrho_{g_j,b_i}(e_{g,j} = 1, e_{b,i} = 1)$ *increases, the probability of both children having an arranged marriage decreases.*

Proof. *This statement follows directly from 14. When* $\varrho_{g_j,b_i}(e_{g,j} = 1, e_{b,i} = 1) < 0$*, parents gain from arranging the marriage of both children, as long as the outside option is sufficiently low. As* $\varrho_{g_j,b_i}(e_{g,j} = 1, e_{b,i} = 1)$ *increases and becomes positive, the dis-utility term of equation 14 dominates and parents prefer to arrange the marriage for only one of their children.*

Figure 14 in the appendix illustrates the results of corollary 3.1, showing that the probability both children (boy and girl) have an arranged marriage decreases as the income correlation between their potential arranged spouses moves from -1 to 1. Column 2 of figure 14 shows, however, that this joint probability still depends on the net difference of returns between boys and girls: the higher the returns for boys, the lower the probability that both children (the boy) have an arranged marriage.[20] Current data does not allow me to test this prediction, I lack of data on the type of marriage for each sibling/child. However, preliminary empirical results suggest that within household there might an endogenous relationship between the gender composition and the probability of having an arranged marriage. In order to explore this prediction for the whole population, it is necessary to incorporate a general equilibrium framework.[21]

[19] The goal of considering the extreme case of a unique insurance partner is to provide a clear intuition on how parents decide how to allocate education and effort. Studying other cases of small networks should deliver a similar intuition, but a more complex analysis might be required.

[20] In all the simulations, I have assumed that boys might get larger returns than girls. This delivers the difference in slopes by gender as shown in the second row of figure 14.

[21] These results are available upon request.

5 Empirical Results

5.1 Test of Full Insurance

The model is based on the assumption that arranged marriage couples are able to smooth consumption over time better than love marriage couples, since on average love marriages should provide less insurance.[22] I use a test of full insurance to examine this assumption. The test is derived by solving the social planner's problem, under the assumption the social planner maximizes a weighted sum of individual household utilities subject to the aggregate income constraint, as follows:

$$Max \sum_{i=1}^{N} \lambda_i \sum_t \beta^t \sum_s \pi_s u_i(c_{ist}) \tag{15}$$
$$s.t. \sum_{i=1}^{N} c_{ist} \leq \sum_{i=1}^{N} y_{ist}$$

The first order conditions of the problem link the marginal utility between two households. Equation 16 shows that the relative consumption of each pair of households will depend on their relative Pareto weights, which are assumed to be time invariant:

$$\frac{\lambda_i}{\lambda_j} = \frac{u'_j(c_{jst})}{u'_i(c_{ist})} \tag{16}$$

The reduced form test depends on the utility function assumed. The typical functional forms used in the literature are constant absolute risk aversion and constant relative risk aversion utility functions. Both utility functions enable us to express the consumption of each individual household only as a function of the aggregate consumption, the Pareto weights and other non-separable variables considered to be relevant (leisure, taste shifters, among others). If a CRRA utility function is assumed, the first order condition can be rewritten as $ln\, c_{jst} = \frac{1}{N}\sum_i ln\, c_{ist} + \frac{1}{\sigma - 1}\left[\frac{1}{N}\sum_i ln\, \lambda_i - ln\, \lambda_j\right]$. This equation links the logarithmic consumption of household j to the aggregate consumption of the economy. It also suggests that consumption of household j will depend on its Pareto weight relative to the average Pareto weight of the economy. In order to eliminate the time invariant Pareto weight component, we might use a fixed effects model or a first differences model:

$$\begin{aligned} &First\, Differences: \; \triangle ln\, c_{ist} = \beta_1 \triangle ln\, \bar{c}_{st} + \beta_2 \triangle ln\, y_{ist} + v_{jst} \\ &Fixed\, Effects: \; ln\, c_{ist} = \alpha_i + \beta_1 ln\, \bar{c}_{st} + \beta_2 ln\, y_{ist} + u_{ist} \end{aligned} \tag{17}$$

Full insurance is not rejected if $\beta_1 = 1$ and $\beta_2 = 0$. In other words, the growth rate of individual consumption moves perfectly with the growth of aggregate consumption, and it does not depend on any other variable, particularly on individual income. The implementation of the test requires panel data on individual or household income and consumption.

[22] Rosenzweig and Stark (1989) study how the variance in consumption is related to marital arrangements. They show that rural households in Southern India use a strategy to diversify risk by marrying their daughters into distant villages and rarely two into the same one, finding that arranged marriages contribute to mitigating the volatility of consumption. India, however, is among the countries where arranged marriages are resilient. Love marriages have recently started to increase in urban areas, growing from 5% to 10%. This feature of the Indian context makes it unsuitable for testing formally whether arranged marriages are indeed used as a commitment device and provide access to insurance for the agents.

5.1.1 Data

I use the Indonesia Family Life Survey to estimate equations 17. The IFLS started in 1993, surveying 7224 households in 13 provinces of Indonesia. It followed these households, in 1997, 2000 and 2007, with a low rate of attrition. The IFLS is the ideal data set to implement this test by type of marriage. It collected detailed information on consumption expenditure (durables and non-durables) and on income (wages, profits from farm and non-farm business, rents from assets, and other non-labor income, among others). The IFLS also collected detailed marital history, in particular, it has information on the type of marriage arrangement for the first marriage of each ever-married individual. Specifically, the IFLS asks respondents to report whether they or their parents chose their spouse.[23]

The sample is restricted to couples where both husband and wife are still in their first marriage in 1993 and remained married through the next two waves (1997 and 2000). I define a household as an arranged marriage household if both spouses self-report having an arranged marriage, and as a love marriage household if both spouses report having chosen their spouses. Finally, I restrict the test to villages and small towns (less than 2500 families) and to those households that did not move outside the village/town between surveys.[24] I assume that the village is the relevant insurance group.

The implementation of the test of full insurance requires information on non-durable consumption expenditure and on non-insured income (Mace, 1991). The consumption component is constructed using monthly information on expenditure on food, utilities, personal toiletries, small household items, recreation and entertainment, transportation, clothing, taxes and rent. Income is calculated using information on labor income, which includes wages, profits from farm and non-farm businesses, and non-labor income (pensions, rents from assets and other bonuses from work). All monetary values are converted to dollars using the 1993 PPP exchange rate.

Table 1 presents the summary statistics for the sample. The final sample contains 1438 households having a love marriage and 410 households having an arranged marriage. There are 313 households reporting a "mixed marriage," where one spouse reports having an arranged marriage and the other reports having a love marriage (in 75% of these marriages the wife reports the arranged marriage). I have excluded these couples since it is unclear who provides insurance in these arrangements. Focusing on 1993, the households having an arranged marriage are almost 8 years older, are slightly larger, and are primarily engaged in agriculture; they are also poorer, their income and expenditure are lower than that of households in love marriages. But in both types of households, expenditure on food represents between 70% and 80% of total non-durable consumption, and the majority of their income comes from labor. Consistent with the theory, households in arranged marriages are more likely to farm and have lower levels of schooling. Interestingly, the "mixed marriages" have an intermediate level of education (both spouses have more schooling than spouses in an arranged marriage, but less than spouses in a love marriage). The last two panels provide suggestive evidence on the extent of informal and formal insurance available for each type of marriage. Transfers to and from other families members are observed for both, though they constitute a larger share of total per capita income

[23]The other data-sets with information on type of marriage (Turkish, Cambodian and Togolese Demographic and Health Surveys, and Vietnam Longitudinal Study), used to derive the stylized facts, do not have information on consumption and/or income. Other data sets with available information on consumption and income do not have information on type of marriage, with the exception of data from India and Bangladesh; however, in both countries most marriages are still arranged, making them unsuitable for the analysis.

[24]Households moving might have lower insurance gains. However, they might also migrate to diversify risk geographically. This is testable in the data.

for arranged marriages in 1997 and 2000.[25] The levels of formal insurance are very low for both groups, but couples in a love marriage have better access to this type of insurance, and in addition, a larger percentage of them also report having savings (and report a higher amount of savings, measured as a percentage of total monthly income) as well as a higher debt/income ratio, suggesting that they have access to other sources of credit.

5.1.2 Results

Table 2 reports the results for the first differences and fixed effects models, where the aggregate consumption is calculated at the village/town level.[26] The results are very similar across all the specifications used. For love marriages, the coefficient on aggregate consumption ranges between 0.342 and 0.474 and it is always statistically significant at a 1% level. The coefficient on individual income is smaller, ranging between 0.00852 and 0.0977, and statistically significant at a 5% level in all cases. These results lead to a rejection of the full insurance hypothesis for couples in love marriages: the joint test of significance presented at the bottom of each panel strongly rejects that the coefficient on aggregate consumption is equal to one and the coefficient on individual income is equal to zero. In contrast, the results for arranged marriages show that income is not statistically significant in the determination of consumption. Furthermore, the coefficient is smaller in magnitude, ranging from -0.00303 to 0.00151. Although, the estimated coefficient on aggregate income is statistically smaller than 1, also leading to the rejection of the full insurance hypothesis, the coefficient on income is always not significant, suggesting that they have access to insurance.

The main concerns in this estimation are measurement error, particularly in income, and omitted variable biases (other variables determine the allocation of consumption, for example, leisure). Appendix tables 9 and 10 perform a series of robustness exercises using a first differences model. The results are robust to dropping the top and bottom 1% of the income distribution and to implementing a robust regression that gives lower weight to observations that might be outliers. They are also robust to using only labor income, adding education and medical expenditure in the non-durable consumption, to using each period separately (1993-1997 and 1997-2000) and to instrumenting for income changes using lags. Overall, all the specifications support the assumption that arranged marriages allow families to share risk —since income does not predict consumption for arranged marriages they appear to be better insured than love marriages; however, the results also suggest that neither love nor arranged marriages have access to full insurance (or that there is a model mis-specification).[27] This is a strong test of insurance — arranged marriage households might not be perfectly insured— rather they might have access to differentially greater levels of insurance. Overall, the results suggest that arranged marriages provide more insurance than love marriages. Next I test the predictions of the model.

---

[25] Although not in 1993, but this year arranged marriage couples reported a higher amount of other non-labor income that includes gifts and arisan winnings, among others.

[26] Some households report losses in profits or zero income in some components. I use two alternative transformations, widely used in the literature, instead of using directly a logarithmic transformation: (i) Inverse sine transformation $asinh(y) = log[y_i + (y_i^2 + 1)^{1/2}]$; and (ii) the neglog transformation $neglog(y) = \begin{cases} log(y+1) & if\ y \geq 0 \\ -log(1-y) & if\ y < 0 \end{cases}$. The results found under either transformation are very similar.

[27] If individuals within groups do not have homogenous preferences, —for instance if they differ in their risk aversion, then the coefficients in this specification are biased (Mazzocco and Saini, 2012). Alternatively, the insurance group may not be correctly defined, also causing bias in the coefficient of aggregate consumption.

5.2 The Indonesian Green Revolution

5.2.1 The Bimas/Inmas Program

The main implication of the model is that the demand for informal insurance determines the level of arranged marriages. I use the introduction of the Green Revolution (GR) in Indonesia to test this prediction. The GR refers to the combined introduction of higher yield rice seeds and improved agricultural techniques that occurred in several developing countries in the second half of the 20th century. Foster and Rosenzweig (1996) show that the GR substantially increased the returns to (primary) schooling in India, possibly because the adoption and implementation of these innovations required learning. The GR also intended to reduce the variance of agricultural output, and thus most likely also reduced the variance of farming households' incomes. The model predicts that both of these changes (lower variance and increased education returns) will result in fewer arranged marriages and higher investment in education. First, I use the timing of the GR in conjunction with variation in the intensity of the program's implementation across regions of Indonesia to identify the effect of the program on the returns to schooling and on the variance of income. In addition, I show that the introduction of this program also increased the income of agricultural households, possibly allowing them to self-insure — a margin not explored in the model but potentially an important mechanism for coping with risk. Then, I use the same variation to identify the effect on marriage arrangements and years of schooling. I discuss the relative effects of the program below.

At the beginning of the 1960s, the Indonesian government started an increasing effort to raise rice production. In 1964/65, they launched in West Java a pilot project known as Demostrasi Massal (Demas) — Mass Demonstration Program— with the goal of testing the suitability of higher yield variety seeds (HYVs) in Indonesian soil. In 1965, the Mass Guidance Program (Bimas) was born. The disappointing results of the first few years led to several changes, and in 1967, the program was divided into Bimas and a new Mass Intensification Program (Inmas). The more intense diffusion of both programs started in Java toward the end of the 1960s and slowly expanded into the other islands; by the mid-1980s around 75% of rice areas were covered under some sort of intensification (Hill, 2000). The program suffered several modifications and additions throughout the years; however, the main components of the program remained unchanged and were composed of: (i) expansion of HYVs; (ii) increased availability of fertilizer and pesticides; (iii) access to credit; and (iv) rehabilitation of irrigation systems (Palmer, 1976). The Bimas/Inmas program successfully increased yields by approximately 65% between 1960/61 (1.4 tons/ha) and 1980/81 (2.3 tons/ha) (Manning, 1995) and eliminated the import of rice, which had reached a maximum during the 1970s.[28]

I collected information for approximately 200 districts on the implementation and outcomes of the Bimas/Inmas program from the 1963 and 1983 agricultural census and other documents on land utilization and agricultural production (appendix table 11). Using the information from the beginning of the 1980s, I first confirm (appendix table 12) that rice areas under Bimas/Inmas used agricultural inputs more intensely: they had a higher percentage of irrigated land; a larger percentage of households used fertilizer and used it more intensely per hectare; and they were more likely to use HYVs. These districts had an average increase of 1.35 tons/ha in rice production, and this increase is explained by a more intense use of agricultural inputs (appendix table 13). The 1963 agricultural census reports information on total sawah land (land available

[28]By 1979 Indonesia was still the largest rice importing country in the world, with annual imports of 2.9 million metric tons. By 1983, Indonesia became a rice self-sufficient country for the first time in its history (Resosudarmo and Yamazaki, 2011)

for the production of wet paddy or rice), total dry land, total land harvested (wet and dry rice), number of farms, average farm size and number of livestock for each agricultural district. Using this information, I show that areas with a higher percentage of sawah land and more farms per hectare (farms/ha) benefited more from the expansion of the Bimas/Inmas program. In other words, both variables and its interaction are good predictors of the implementation of the Bimas/Inmas program (appendix tables 14 and 15). The reduced form analysis, using the percentage of sawah land and farms/ha instead of the intensity of the program, indicates that these areas had an average increase of 1.28 tons/ha in rice yield.[29]

5.2.2 The Effects of the Green Revolution on Labor Market Outcomes

To document the effects of the GR, I use data from the 1976 and 1995 Population Census, and the 1980, 1981, 1982, 1984 and 1987 socioeconomic surveys (SUSENAS).[30] I also confirm these results with the IFLS 1993 sample.[31] Specifically I estimate the following Mincerian wage equation:

$$ln\, wage_{i,p,t} = \beta_o + \pi_t + \delta_d + \beta_1 s_{i,d,t} + \beta_2 (s_{i,d,t} * bimas_{d,1982}) + X_{i,d,t} + \varepsilon_{i,d,t} \qquad (18)$$

where the log wages of individual i in district d at time t are a function of (survey) year fixed effects (π_t), district-of-residence fixed effects (δ_d), the years of education ($s_{i,d,t}$) and other individual variables, $X_{i,c,t}$ (gender, age, age squared and a dummy for rural residence). The variable $bimas_{d,1982}$ measures the intensity of the program as the percentage of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83. The coefficient of interest is β_2, and it tests whether the returns to school are larger in areas where the GR had a larger impact.

The identification relies on variation in the intensity of the treatment across districts and cohorts, and assumes that there are no unobserved determinants of wages (in particular of the returns to school) that are correlated with the timing and intensity of the program. Outcomes before the mid-1960s/early-1970s should not have been affected by the program; in addition, the intensity of the treatment must have varied across younger generations, i.e., some individuals were exposed to the program for a longer period of time (parents and children reaped the benefits earlier relative to other areas receiving the program at a later period). The main threats to identification are differential trends across areas and adoption of other programs at the district/cohort level. In order to mitigate these potential concerns of endogeneity (less conservative areas, with higher wages, adopting the HYVs earlier and having a faster transition towards love marriages), I additionally instrument $bimas_{d,1982}$, the intensity of the GR, with the percentage of sawah land in 1963, the number of farms per hectare in the same year and their interaction (taken from the 1963 agricultural census).

Another important issue to consider is that the GR coincided with the introduction of other programs. Indonesia experienced an intense school construction program between 1973 and 1978, exploited by Duflo (2001), who shows that it successfully increased the education of children across districts and cohorts. During the same period, the government also expanded sanitation and water supply across areas. Figure 12 plots the intensity of the Bimas/Inmas program (% of sawah land under intensification) against the intensity of the

[29] Appendix tables 14 and 15. The reduced form refers to the following equation: $yieldrice_{d,1980} = \beta_o + \beta_1 \% sawahland_{d,1963} + \beta_2\, (farms/ha)_{d,1963} + \beta_3 \%\, sawah\, land_{d,1963} * (farms/ha)_{d,1963} + \varepsilon_{d,1980}$

[30] The 1976 Census is the first survey containing information on any type of income, collecting for the first time information on wages (only for employees). The 1980, 1981 and 1982 SUSENAS and the 1995 Census collected information on wages and educational attainment. The 1981, 1982, 1984 and 1987 SUSENAS surveys collected detailed information on income at the household level from all sources (including profits from agricultural and non-agricultural business).

[31] Appendix 8.5 and tables 16 and 17 present a more detailed description of the data construction and report the summary statistics for the estimation samples.

school construction program (# schools for 1000 children), and against the allocation of water and sanitation programs (per capita) by district. In both cases, there seems to be a slightly negative correlation between each pair of variables, mostly driven by one district, suggesting the omitted variables bias (OVB) is potentially unimportant—I add these controls as a check.

Panel A of table 3 reports the results for the pooled sample of SUSENAS and Census. The coefficients show that each additional year of schooling increases the returns to schooling between 5.4% and 5.9%, while individuals residing in areas exposed to the average intensity of the Bimas/Inmas program have a statistically additional return of 2.1% to 2.6% for each extra year of schooling. The table also shows the reduced form effect of the 1963 agricultural characteristics on log wages and the IV results of instrumenting $bimas_{d,1982}$ with them. In both cases, the effect of the GR is larger, raging between 3.3% to 4.7%. Panel B presents the results for the 1993 cross-section of IFLS. The results are slightly larger for this sample; each year of education increases wages by 7.4% to 8%, and an additional 3.6% to 3.9% for individuals living in districts with average exposure to the Bimas/Inmas program. Table 18 in the appendix adds controls for the school construction program and the water and sanitation projects; the results of the direct effect of Bimas/Inmas are slightly smaller but still significant at a 1% level (columns 1 and 2) suggesting that the effect of the Bimas/Inmas program is not capturing an additional effect of these programs. Reassuringly, the IV results do not change.

Table 4 presents the results of the joint estimation of the effect of the program on the mean and variance of income using the following specification

$$ln\, income_{i,d,t} = \beta_o + \pi_t + \eta_p + \beta_1 bimas_{d,1982} + \beta_2 rural_{i,d,t} + \sum_{2}^{4}(\pi_t * bimas_{d,1982})\beta_{3,t} + \varepsilon_{i,d,t} \qquad (19)$$

where $ln\, income_{i,d,t}$ is the logarithm of the per capita income of household i, residing in district d at survey time t, π_t and $bimas_{d,1982}$ are defined as in equation 18, η_p are a set of province fixed effects, $rural_{i,d,t}$ is dummy variable for rural residence, and $(\pi_t * bimas_{d,1982})$ captures differences of the Bimas/Inmas program for each survey year. The specification jointly estimates the effect on the mean and variance of (log per capita) income using a maximum likelihood model that assumes normally distributed errors.

Panel A contains the results for the SUSENAS sample using only agricultural households. The first two columns present the results for specification 19. In columns 3 and 4, I add a dummy variable for two islands (Sulawesi and Kalimantan) interacted with $bimas_{d,1982}$, π_t and $\sum_{2}^{4}(\pi_t * bimas_{d,1982})$ because these islands received the intensification program late (towards the end of the 1970s and beginning of the 1980s); the districts on these islands had less time to show a perceivable change in production.[32] The coefficients in columns 1 and 3 show an average income increase between 6.7% and 8.2%.[33] The results in column 2 would suggest an average increase in the income variance of 2.9%; however, this coefficient masks the fact that some districts (the outer islands) received the program much later and had not reaped the benefits by the time of

[32] The specification used is: $ln\, income_{i,d,t} = \beta_o + \pi_t + \eta_p + \beta_1 bimas_{d,1982} + \beta_2 rural_{i,d,t} + \sum_{2}^{4}(\pi_t * bimas_{d,1982})\beta_{3,t} + \beta_4(bimas_{d,1982} * islands_{i,d,t}) + \sum_{2}^{4}(\pi_t * islands_{i,d,t})\beta_{5,t} + \sum_{2}^{4}(\pi_t * islands_{i,d,t} * (bimas_{d,1982})\beta_{6,t} + \varepsilon_{i,c,t}$

[33] The coefficients of the interaction terms, $\beta_{3,t}$, are not shown in the tables but they do indicate that the program had a different impact by year, increasing income further in 1982 and 1987, and slightly less in 1984, although always having a positive effect relative to areas with lower treatment intensity.

the surveys (as also suggested by the slightly lower coefficient in column 1). Once I account for this fact in column 4, the estimated coefficient shows a decrease in variance of 8.2%.

The reduced form effect of the 1963 agricultural characteristics in columns 7 and 8 (once I control for the two islands) indicate a smaller gain in income and much larger decrease in variance, 0.5% and 46% respectively. Panel B repeats the same analysis for the sample of agricultural households of the 1993 IFLS cross-section; the results are much larger in magnitude, but the signs and conclusions remain unchanged (both for the OLS and the reduced form analysis). Table 19 in the appendix adds controls for the other programs implemented during the same period (school construction, water and sanitation supply). The coefficients do not change significantly and the signs are the same as before.

5.2.3 Effect on Arranged Marriages and Education

The results from returns to education indicate that the program caused an increase in the outside option and, at the same time, decreased the size of the expected shock for agricultural households.[34] The model predicts that these changes should increase the incentives to invest in education (increasing years of schooling) and speed up the transition toward love marriages (decreasing the probability of having an arranged marriage). I use the following specification in order to test the effect on arranged marriages:

$$AM_{i,d,c} = \beta_o + \gamma_c + \delta_d + \sum_c (\gamma_c * bimas_{d,1982})\beta_{1,c} + \sum_c (\gamma_c * P_{d,61})\beta_{2,c} + \beta_3 female_{i,d,c} + \sum_c (\gamma_c * X_d)\beta_{4,c} + \varepsilon_{i,d,c} \quad (20)$$

where $bimas_{d,1982}$ is defined as before, δ_d are district-of-birth fixed effects, γ_c are cohort-of-birth fixed effects, $P_{d,61}$ is total population residing in district d in 1961, $female_{i,c,d}$ is a dummy that takes the value of one if individual i, in cohort c and in district d is a woman; and X_d includes a set of district level controls from the 1971 census, the intensity of the school construction program from the mid-1970s and controls for the expansion of water and sanitation supply programs during the same period. The coefficients of interest are $\beta_{1,c}$, which represent the effect of the Bimas/Inmas program on the probability that individual i, born in district d and in ten-year birth cohort c (year of birth: 1932<, 1933-42, 1943-52, 1953-62 and 1963>) has in an arranged marriage. If the program was exogenous and unanticipated, we should not observe any effect for individuals married before 1970 (before the GR took place). The average age at marriage is approximately 18; thus, we should expect an effect for individuals born after 1953. The specific timing should depend on how long it took for the program to be implemented and to deliver successful outcomes that are perceived as permanent by households.

The identification strategy is ilustrated in figures 8 and 9, and in table 21 in the appendix. The first figure plots the density of the residuals of the treatment variable after controlling for the total population by district in 1961 and province fixed effects. It then adds controls for district characteristics of 1971. And finally, it controls for the intensity of the school construction, and the water and sanitation supply program. The density does not change as I add information at the district level suggesting that the intensity of the treatment in the early 1980s is not correlted with district characteristics in 1970s. The results are also shown in appendix table 20 where I show that I cannot reject that the 1971 district characteristics and the intensity of other programs are jointly not statistically significant determining the intensity of the treatment in 1982.

[34]Additionally, the increase in income might allow households to self-insure, which it is not captured by my model but might be another potential channel. Also, it should be noted that I cannot separate the effects of increasing returns to education, increasing income and decreasing income variance; however, all these changes should lead to the reduction of arranged marriages according to the theoretical framework.

In figure 9 I divided the treatment variable into the four quartiles of the distribution of sawah land covered by any intensitifcation program. The figure plots the mean AM by cohort of birth and treatment intensity, where treatment one corresponds to the lowest treatment (the districts with the lowest percentage of sawah land covered by the GR). The figure shows that cohorts 1933-42, 1943-52 and 1953-62 follow the same trends in the four treatment intensity areas. It also shows the that identification is obtained from the last cohort by comparing areas more intensily treated against areas with low treatment intensity (after differencing out the common pre-trends). The figure also suggests that the oldest cohort (born before 1933) might not be the proper comparison group. I show results using all cohorts and using only the 4 youngest cohorts. Finally table 21 shows the difference-in-difference raw results in a table with the four treatments. Panel A show the results of the **placebo tests** using the first three cohorts (not exposed to the green revolution), the difference-in-difference specification shows no effect for them. Panel B presents the experiment of interest, showing a large effect for this cohort and slighly increasing with the intensity of the treatment. These results are formally explored using equation 20.

Panel A of table 5 presents the summary statistics for the final sample used. I matched 9,068 individuals to their district of birth using retrospective information on migration. Each cohort has between 1000 and 2500 observations. As already summarized in section 3, arranged marriages have decreased in each generation and education has increased. Interestingly, the percentage of the population currently residing in a rural area has remained relatively stable at approximately 60%. Panel B shows the results of the reduced form effect of the Bimas/Inmas program on arranged marriages using 4 cohorts (columns 1 to 4) and 5 cohorts (columns 5 to 8) estimating specification 20. Column 1 (5) shows the base results, column 2 (6) adds controls at the district level interacted with cohort fixed effects, column 3 (7) adds controls for the school construction program as used by Duflo (2001), finally column 4 (8) adds controls for the expansion of water and sanitation supply. The results support the predictions of the model. The only statistically significant coefficient is for the youngest cohort (born after 1963). Adding controls at the district level does not change the size of the effect, while adding controls for the school construction and the water and sanitation supply programs reduces the magnitude of the coefficient, but still shows a large decline in arranged marriages for the cohorts exposed to the GR (although imprecisely estimated). The results imply that the areas and cohorts exposed to the mean treatment intensity had a 9 to 20 percentage points reduction in the probability of having an arranged marriage, a decline on the order of 30 to 66 percent relative to the sample mean.

Accoding to the theoretical framework, increases in returns to education and the reduction in the net benefits of the insurance of arranged marriages should lead to an increase in years of schooling for the youngest cohort exposed to the program (assuming that the opportunity cost of child labor remains smaller relative to the increase in returns). To test this prediction, I estimate the following equation:

$$Yrs\,sch_{i,d,c} = \beta_o + \gamma_c + \delta_d + \sum_c (\gamma_c * bimas_{d,1982})\beta_{1,c} + \sum_c (\gamma_c * P_{d,61})\beta_{2,c} + \beta_3 female_{i,d,c} + \sum_c (\gamma_c * X_d)\beta_{4,c} + \varepsilon_{i,d,c} \quad (21)$$

where all the variables have the same definition as in specification 20 and $Yrs\,sch_{i,d,c}$ represents the years of schooling of individual i, born in district d and beloging to cohort c. The coefficients of interest are $\beta_{1,c}$, which represent the effect of the Bimas/Inmas program on schooling. The Bimas/Inmas program should have affected the education of children of school age by the time that it delivered benefits. Since the program started in the mid-1960s and slowly expanded, average education should start increasing for children born after 1963 (of school age by early 1970).

The results are reported in table 6. The first column presents the results for equation 21. Column 2 (4) adds controls at the district level interacted with cohort fixed effects, and column 3 (6) adds controls for the school construction program and for the expansion of water and sanitation supply. The coefficients in column 1 suggest that all cohorts living in areas with higher intensity of the program have more education; however, the results show a 0.3-0.5 years increase for the youngest cohort relative to the two previous generations. Controlling for the school construction program reduces the magnitude of the coefficients for all cohorts but the relative increase remains the same. The school construction program targeted areas with low enrollment in elementary school. The reduction in the coefficients suggests that the previous results may be explained by differences in supply of schooling across districts.

The results on the probability of having an arranged marriage and years of schooling support the main predictions of the model. As the net insurance benefits of arranged marriages decline relative to the outside option, parents invest in more education for their children, and children switch faster to love marriages. I next present several robustness exercises and then the test of the divorce extension of the base model.

5.2.4 Robustness Checks

I perform several robustness exercises that deliver similar results to those of the previous section. In the case of arranged marriages, I first instrument $bimas_{d,1982}$ with agricultural characteristics of 1963: the percentage of sawah land, the number of farms per hectare and their interaction (taken from the 1963 agricultural census). I then use splines defined as the four quartiles of the treatment intensity.[35] I also allow for other non-linear effects by introducing a quadratic term of the treatment variable. Finally, I use an alternative defition of treatment. I redefine treatment as the percentage of total agricultural land covered by any intesification program.

The results of the first exercise are presented in table 22 in the appendix. I present the results using 4 and 5 cohorts, adding sequentially controls for the introduction of other program.[36] The base results are larger in magnitude than the OLS results presented in the previous section and they imply a decline of 25 to 42 percetange points. The magnitude of the decline decreases as I add controls for the other programs implemented during the 1970s. However, the results still support the hypothesis of my model.

Table 23 shows the results using splines for the Bimas/Inmas intensity program (columns 1 to 4) and for the alternative definition of treatment (columns 5 to 6), estimating the following equation:

$$AM_{i,d,c} = \beta_o + \gamma_c + \delta_d + \sum_{c=2}\sum_{j=2}(\gamma_c * T_j)\beta_{1,c,j} + \sum_{c}(\gamma_c * P_{d,61})\beta_{2,c} + \beta_3 female_{i,d,c} + \sum_{c}(\gamma_c * X_d)\beta_{4,c} + \varepsilon_{i,d,c} \quad (22)$$

where T_j is a dummy variable for each quartile of the treatment distribution, the omitted category corresponds to the first quartile. The rest of the variables are defined as in equation 20. I present the baseline results where I only control population in 1961 and the results with the full set of controls (district level characteristics in 1971, the intensity of the school construction program and the expansion of water and sanitation supply program).[37] The results of all the specifications still show that cohorts affected by a higher intensity of the GR transitioned faster to love marriages. The point estimates of my preferred specification

[35] Treatment intensity refers to the percentage of sawah land cover by any intensification program

[36] I do not include controls at the district level in 1971. Since my instrumental variables are taken from the 1963 agricultural census, they might have likely influenced the district characteristics in 1971.

[37] The other results are available upon request.

using Bimas/Inmas (colums 2 and 4 which have the full set of controls) imply a reduction of 7.4 to 10.6 percetange points for the case where the base is the cohort born in 1933-43; and a reduction of 17.9 to 19.7 percentage points for the 5 cohorts case. Columns 6 and 8 use the alternative definition of treatment. In the case of the 4 cohorts specification, the reduction implied by the results is between 10.1 and 16.3 percentage points; while for the 5 cohorts case, the decline is between 15.5 and 20.4 percentage points. All these results are within the same range found in the previous section.

Table 24 allows for concavity or convexity in the effect of the treatment by including a quadratic term:

$$AM_{i,d,c} = \beta_o + \gamma_c + \delta_d + \sum_c (\gamma_c * Treatment_{d,82})\beta_{1,c} + \sum_c (\gamma_c * Treatment^2_{d,82})\beta_{2,c} + \beta_3 female_{i,d,c} + \sum_c (\gamma_c * X_d)\beta_{4,c} + \varepsilon_{i,c,d} \quad (23)$$

Similar to table 23, I present the baseline results where I only control for the district population in 1961 and the results with the full set of controls for the both treatment variables. At the bottom of the table, I present the total effect and its standard errores calculated at the mean of the treatment intensity.[38] The results are again within the range found in the previous exercises, the marginal effect evaluated at the mean of the treatment intensity implies a decline in arranged marriages between 8 to 30 percentage points.

Finally, I use the the 1995 inter-census survey to study the effect on investment in education as a robustness exercise. I estimate equation 21 using larger samples, 327,404 individuals when I use only 4 cohorts, and 360,383 when I use the 5 cohorts. The results are shown in table 25, where I also present the results of instrumenting my treatment intensity with the 1963 agricultural characteristics. The point estimates are slightly different from the results using IFLS (the mean education in the the inter-census survey is higher than the mean education in IFLS). However, the implied increased caused by the GR for the cohorts exposed is similar to the previous results. The estimates imply an increase between 0.22 to 0.41 years of schooling relative to the two previous cohorts (using either 4 or 5 cohorts and evaluated at the mean treatment intensity). The IV results are larger in magnitude, suggesting an increase of an additional 0.6 to 0.9 years of schooling for the youngest generation relative to the previous two cohorts. Although not shown in tables, similar results are found using splines, adding a quadratic term for the treatment variable and using the alternative definition of treatment intensity.[39]

5.2.5 Additional Prediction: Divorce

Indonesia and other countries from Southeast Asia have traditionally had a low cost of divorce.[40] Prior to its independence and until 1974, Indonesia had a plural marriage law system. The population was divided into five groups (Muslims, Christians, Chinese, Europeans and natives) and each of them had its own law.[41] However, the large majority of the population was subject to the unwritten customary (adat) law in combination with the Muslim Family Law. In 1974, the government approved the Family Law Bill, which provided a unified framework (keeping a separation between Muslims and other religions). The law effectively increased the cost of divorce for all individuals through enforcing the registration of the marriage and requiring court approval before any divorce was effective. This increase in the cost of divorce occurred during the same period when the demand for informal insurance fell. Younger cohorts faced an increasing drop in the net insurance benefits as risk profiles shifted, the cost of informal insurance increased, and the outside option

[38] $Total\,effect{=}\beta_{1,c} + \beta_{2,c} * 2 * (Treatmet)$

[39] Results available upon request

[40] Jones (1981, 1997) presents and discusses evidence on the trends of divorce in Indonesia, Malaysia and Singapore since 1950.

[41] For Muslims, Islamic Law applied; for indigenous Indonesians and some "Foreign Orientals," their customary law applied; for Indonesian Christians the marriage ordinance applied; and for Chinese and European the Civil Code applied.

raised. Holding constant the cost of divorce, the model predicts that we should observe an increasing divorce rate among arranged marriage couples; however, since at the same time the cost of divorced went up, the effect might be smaller. To test this prediction, I use the following specification:

$$D_{ipc} = \beta_0 + \gamma_c + \eta_p + \beta_1 AM_{ipc} + \sum_c (\gamma_c * AM_{ipc})\beta_{2,c} + \sum_p (\eta_p * AM_{ipc})\beta_{3,p} + \beta_4 female_{ipc} + \beta_5 duration_{ipc} + \varepsilon_{ipc} \quad (24)$$

where D_{ipc} takes the value of one if individual i in province p and cohort c divorced her first spouse, γ_c are cohort fixed effects, η_p are province fixed effects, AM_{ipc} is a dummy variable taking the value of one if individual i in province p and cohort c had an arranged marriage, $female_{ipc}$ takes the value of one for females, and $duration_{ipc}$ controls for the number of years that the first marriage lasted. The coefficients of interest are $\beta_{2,c}$ and the coefficients on γ_c . The former captures the probability of divorce among couples in arranged marriages for each cohort (the omitted category is the oldest cohort, born before 1933), while γ_c capture the divorce probability among couples in love marriages. Figure 10 plots the coefficients of a linear probability model. The results support the prediction of increasing divorce rates among arranged marriage couples for younger cohorts, despite the increased legal cost of divorce. Interestingly, they also show a decreasing divorce rate for individuals in a self-choice marriage even prior to the 1974 family law change. These results might be the outcome of better matching in the love marriage markets and possibly other changes that lowered their threshold for divorce. For example, individuals might have improved their search process in love marriage markets finding mates that provide them insurance,$\varrho_{ks}\,(e,\, I(L = 1))$, for new types of risk, σ^2_{δ}.

I additionally use the exogenous variation generated by the Green Revolution to provide more convincing evidence on the declining insurance benefits of arranged marriages. I estimate the following equation by type of marriage using the first three cohorts of my sample (couples mostly married before mid-1960s):[42]

$$\begin{aligned} D_{idc} = {} & \beta_0 + \gamma_c + \delta_d + \sum_2^3 (\gamma_c * bimas_{d,1982})\beta_{1,c} + \sum_2^3 (\gamma_c * ymb65_{idc})\beta_{2,c} + \\ & \sum_2^3 (\gamma_c * ymb65_{idc} * bimas_{d,1982})\beta_{3,c} + \beta_4 female_{idc} + \beta_5 duration_{idc} + \varepsilon_{idc} \end{aligned} \quad (25)$$

where γ_c, $female_{idc}$ and $duration_{idc}$ are defined as in equations 24; δ_d are district-of-birth fixed effects and $bimas_{d,1982}$ measures the intensity of the program as the percentage of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83.[43] Finally, $ymb65_{idc}$ is dummy variable that takes the value of one if the individual was married before 1965. The coefficients of interest are $\beta_{3,c}$, that capture the effect of the GR on individuals married before the Bimas/Inmas was implemented and allowing the effect to vary by cohort.[44] The results in table 7 show that individuals having an arranged marriage and exposed to the GR indeed have an increasing probability of divorce. The results from column 1 suggest that the cohort born before 1933 and exposed to the Bimas/Inmas program has a 18% lower probability of divorcing; however, this probability increases to 25% for the cohort born between 1933 and 1942, and to 29%

[42]I am concerned that the introduction of the GR changed selection into marriage for both arranged and love marriages. Thus, I focus on couples married before it initiated.

[43]I am using district of birth to assign the intensity of the Green Revolution. It might not be the proper measure if individuals in arranged marriages and love marriages have different migration patterns.

[44]Cohorts married before 1965 experienced different historical and economic changes, which might affect differentially the selection into marriage. For instance, couples married before 1950 (year of birth before 1933) were formed during the war period.

for the cohort born between 1943 and 1952. The results from column 2, which uses data from individuals having a love marriage as a placebo test, suggest that the GR did not change significantly the probability of divorce for this group.

5.3 Other countries: Divorce in Turkey

Finally, I study divorce trends in Turkey using the 2003 and 2008 cross sections of the Turkish Demographic and Health Survey (TDHS). The TDHS collects information on ever-married women whose ages are between 15 and 49; I focus on women born before 1983, leaving a total of 7321 from the 2003 cross-section and 5884 from the 2008 TDHS. In contrast with Indonesia, Turkey has had a low rate of divorce during the last century (around 2% of ever-married women have had a divorce) and women have been granted equal rights to men almost since its foundation (after the foundation of the Turkish Republic in 1923, the family law banned polygamy, and divorce and inheritance rights were made equal for men and women; and in the 1930s women acquired full political rights, including the right to elect and be elected locally and nationwide). Figure 11 plots the (raw) percentage of women in an arranged marriage by year of marriage. It shows decades of accelerated decline (1960 to 1970 and mid-1990s to late-1990s) and other phases with a slower and smoother transition (1970s, 1980s and 2000s).

I use the following specification to explore changes in divorce rates by cohort and type of marriage:

$$D_{idct} = \beta_0 + \gamma_c + \delta_d + \beta_1 AM_{idct} + \sum_c (\gamma_c * AM_{idct})\beta_{2,c} + \beta_3 duration_{idct} + \pi_t + \varepsilon_{idct} \tag{26}$$

where where D_{idct} takes the value of one if woman i in district d, cohort c and survey year t divorced her first spouse, γ_c are cohort fixed effects, δ_d are district-of-residence fixed effects, AM_{idct} is a dummy variable taking the value of one if the woman had an arranged marriage, $duration_{ipc}$ controls for the number of years that the first marriage lasted and π_t is a survey-year fixed effect. The coefficients of interest are β_1, $\beta_{2,c}$ and the coefficients on γ_c. Panel A of table 8 presents summary statistics for the sample used; women in love marriages are slightly younger, have approximately 3 years more of education and are more likely to live in urban areas. Panel B shows the results of equation 26, showing that women having an arranged marriage are 1% more likely to have had a divorce (columns 1 and 2) than women having a love marriage. This effect is statistically significant at a 1% level. The next two columns (3 and 4) suggest that this result is mostly driven by women born between 1964-1973 and 1974-1983, who are 0.9% and 0.62% more likely to have divorced, respectively. These results are especially interesting when analyzed jointly with the patterns of figure 11. Women born between 1964 and 1973 married approximately in 1983-1992 (average age at marriage is 19 years old), a period where the decline in arranged marriage was slow; but they divorced during the mid-1990s (average duration of marriage is 7.5 years), a period of rapid decline in arranged marriages. The theoretical framework would suggest that this period experienced a rapid decline on the insurance benefits of arranged marriages, thus leading to a larger rate of divorce as found in the empirical analysis. In contrast, women born between 1974 and 1983, married between 1992-2003. This period (1992-2003) witnessed another accelerated rate of decline in arranged marriages (possibly also changing selection into marriage). This last cohort divorced during the 2000s, again a period of slower decline in arranged marriages, which possibly resulted in the slightly lower probability of divorce for this cohort.

6 Discussion

This paper documents a transition in marital arrangements in several regions of Asia and Africa: arranged marriages are disappearing. To understand and explain the causes behind this change, I document the factors associated with this transition. Overall, arranged marriages have declined as the returns to school have increased and countries have moved away from agriculture, becoming more urban. Based on these observations and on previous literature, I propose a simple model of marital formation to understand how these economic changes have affected the incentives of households for entering into an arranged marriage.

The main assumption of the model is that arranged marriages provide a form of informal insurance that other marriages do not. In the model, parents invest in the education of their child and exert an effort to find her an insurance partner (the arranged marriage). The child receives the returns to education, transfers a share back to her parents and decides to accept the arranged marriage or not. The main implication of the model is that a net reduction of the insurance benefits relative to the outside option leads to the disappearance of arranged marriages (both children and parents face an increasingly costly trade-off). I extend the model to allow for the possibility of divorce. I show that there will be a higher divorce rate among arranged marriage couples as their insurance benefits disappear.

I provide empirical evidence in support of the theoretical framework. Households in arranged marriages are better insured than those in love marriages: changes in income do not predict changes in consumption among arranged marriage couples, but they do for love marriage couples. Then, I use the introduction of the Indonesian Green Revolution as an exogenous technological shock to the distribution of earnings that increased the returns to schooling, lowered the variance of income, and increased the level of income of agricultural households that traditionally married by parental arrangement. As predicted by the model, the intensity of the GR accelerated the transition toward love marriages and increased the investment in education for the cohorts (and areas) more exposed to the technological innovation. Finally, using data from Indonesia and Turkey, I also show that arranged marriage couples have had an increasing divorce rate consistent with the declining insurance benefit that these marriages provide.

Overall, the results presented in this paper are consistent with a relative decrease in the insurance value of arranged marriages. The net benefits of this type of informal insurance arrangement have decreased relative to the (unconstrained) returns outside of the social network. The transition, however, might also be consistent with two alternative (or complementary) explanations. First, in an alternative model, children might have experienced an increase in their bargaining power as older generations lose control over the resources of the economy. Being excluded from land inheritance or denied access to traditional occupations might no longer be a sufficiently severe punishment to influence the children's decisions. Second, there has been an increasing media penetration highly correlated with economic growth that might lead to cultural changes or westernization of these regions. Both of these alternatives are also compatible with the decrease in the insurance motive I propose (potentially increasing the speed of the marital transition). Moreover, these explanations cannot fully account for all the patterns, in particular, for early changes in some countries (especially for more rural regions where media might have arrived only in recent decades).[45] However, I

[45] The first alternative explanation might be incorporated in a framework where parents use their children to increase their ties to the community (create alliances or conserve social status) in order to have access to other benefits (which might be considered as a broader insurance motive). As children increase their relative bargaining power, their tastes have a bigger role in the determination of the final choices. The movement away from arranged marriages would suggest that children obtain lower utility from this particular type of insurance; however, it does not imply that they will not have access to other forms of social

cannot disentangle them with the current data.

This paper leaves open several questions for future research. First, there is the puzzle of South Asia. This region, and India in particular, has experienced high economic growth and several economic changes in the last decades. Foster and Rosenzweig (1996) and Munshi and Rosenzweig (2009) show that the Green Revolution increased the returns to schooling and the income of agricultural households. Despite these changes, arranged marriages remain resilient. The low mobility in marital arrangements is consistent with the fact that marriage takes place within the sub-caste or jati, constraining geographic and social mobility (documented by Munshi and Rosenzweig (2009) among others). The jati acts as an insurance network, allowing households to spread across regions and reduce the impact of negative shocks. Mobarak and Rosenzweig (2012) and Munshi and Rosenzweig (2006), among others, show that the jati plays an important role in business investments, in employment (in rural and urban areas), and in risk sharing. Their networks span large regions and maintain rules of strict marital endogamy. In this context, the second extension of the model may help explain why South Asia is resilient to the economic changes. As households have more children, they benefit from insurance only if they are able to diversify their marriages across different areas. Access to large and efficient social networks guarantees that parents are able to find good insurance partners for all or most of their children. Moreover, if the networks span large geographic areas, they also might reduce the value of the outside option. Then constrained returns to schooling are closer to unconstrained returns under these conditions. Larger net benefits and a smaller outside option lower the impact of the economic changes, consistent with the evidence presented in this paper. The marital transition, however, has started in urban areas (love marriages have increased from 5% to 10%) where the outside option is rapidly increasing. Munshi and Rosenzweig (2009) also document a similar pattern in Mumbai, where marriage outside the jati has grown from 2% to 12%. More research is needed to show that economic agents indeed have larger insurance gains and a lower outside option in South Asia.

Another open question is the determination of demand and supply in the marriage markets as countries move from arranged to love marriages. The second theoretical extension of the model, where I study the case of households with two children and varying gender composition, shows that if children are heterogeneous (in price and returns to schooling) and parents belong to a small social network, they have incentives to arrange the marriage only of the child with the lowest expected return in the labor market (and satisfy their insurance needs). If women are considered to have lower returns than men (Strauss and Thomas, 1996; Behrman, 1997), then parents will prefer to arranged the marriage of their daughters. Preliminary results using data from Indonesia, Turkey and Vietnam, show that gender sibling composition is indeed associated with the probability of having an arranged marriage (even after instrumenting this measure).[46] However, in equilibrium, households should supply the same number of boys. It is not clear then how the marriage markets would reach the equilibrium. More research in needed in this direction.

Finally, recent literature in development economics has shown important welfare losses for some individuals living in arranged marriage societies (South Asia). Field and Ambrus (2008) show that women in Bangladesh attain less schooling as a result of social and financial pressure to marry young. Vogl (2013) shows that arranged marriage cultivates rivalry among sisters in South Asia. During spousal search, parents with multiple

risk-sharing arrangements. The second alternative explanation might also be considered as complementary if media penetration sped up the transition across areas within a country (experiencing high economic growth) by disseminating information on new employment opportunities (at the same time that tastes of parents and children shifted).

[46]Available upon request.

daughters reduce the reservation quality for an older daughter's groom, rushing her marriage to allow sufficient time to marry off her younger sisters. Younger sisters cause earlier exit from school, lower literacy, a match to a husband with less education, and lower adult economic status. This paper further shows that arranged marriages are unstable in some of these regions (have a high divorce rate), leaving open the question of the welfare cost for their offspring. Moreover, the theoretical paper by Edlund and Lagerlof (2006) also suggests that love marriages might reinforce the process of economic growth. Thus, the transition to self-choice marriages might have important effects on the distribution of welfare. Answering the question of what drives the transition is the first step toward exploring who gains and who loses in this "love" revolution.

References

[1] Ackerberg, Daniel A. and Maristella Botticini. 2002. "Endogenous Matching and the Empirical Determinants of Contract Form." *Journal of Political Economy*, 110(3): 564-591.

[2] Anderson, Siwan. 2007. "The Economics of Dowry and Brideprice." *Journal of Economic Perspectives.* 21(4): 151-174.

[3] Banerjee, Abhijit V., and Esther Duflo. 2007. "The Economic Lives of the Poor." *Journal of Economic Perspectives*, 21(1): 141-168.

[4] Barcellos, Silvia, Leandro Carvalho and Adriana Lleras-Muney. Forthcoming. "Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?." *AEJ: Applied Economics.*

[5] Behrman, Jere. 1997. "Intrahousehold Distribution and the Family," in Mark Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics.* Amsterdam: Elsevier Science.

[6] Becker, Gary and Isaac Ehrlich. 1972. "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy.* 80 (4): 623-48.

[7] Becker, Gary.1973. "A Theory of Marriage: Part I," *Journal of Political Economy* 81 (4): 813-46.

[8] Becker, Gary. 1974. "A Theory of Marriage, Part II," *Journal of Political Economy* 82 (no. 2, part 2): S11-S26.

[9] Becker, Gary. 1991. "A Treatise on the Family," Harvard University Press, expanded edition.

[10] Browning, Martin, Pierre-Andre Chiappori and Yoram Weiss. 2014. *Economics of the Family (Cambridge Surveys of Economic Literature).* Cambridge University Press: Cambridge.

[11] Buunk, Abraham, Justin H. Park and Shelli L. Dubbs. 2008. "Parent-Offspring Conflict in Mate Preferences." *Review of General Psychology*, 12(1): 47-62.

[12] Calomiris, Charles W. and Indira Rajaraman. 1998. "The role of ROSCAs: lumpy durables or event insurance? " *Journal of Development Economics*, 56: 207–216

[13] Chang, Kyung-Sup. 1997. "The Neo-Confucian right and family politics in South Korea: The nuclear family as an ideological construct." *Economy and Society*, 26(1): 22-40.

[14] Coate, Steven, and Martin Ravallion. 1993. "Reciprocity without Commitment: Characterization and Performance of Informal Insurance Arrangements," *Journal of Development Economics*, 40: 1-24.

[15] Cochrane, John H. 1991. "A Simple Test of Consumption Insurance." *Journal of Political Economy* 99(5): 957–76.

[16] Cole, Shawn , Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery. 2013. "Barriers to Household Risk Management: Evidence from India," *American Economic Journal: Applied Economics*, 5(1): 104–135.

[17] de la Briere, Bénédicte, Elisabeth Sadoulet, Alain de Janvry and Sylvie Lambert. 2002. "The roles of destination, gender, and household composition in explaining remittances: an analysis for the Dominican Sierra." *Journal of Development Economics*, 68: 309– 328.

[18] Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795-813.

[19] Edlund, Lena and Nils-Petter Lagerlof. 2006. "Individual vs. Parental Consent in Marriage: Implications for Intra-household Resource Allocation and Growth," *American Economic Review P&P*, 96(2): 304-307.

[20] Fafchamps, Marcel, and Agnes R. Quisumbing. 2007. "Household Formation and Marriage Markets in Rural Areas." In T. Paul Schultz and John Strauss, eds., *Handbook of Development Economics*, 4. Amsterdam: Elsevier.

[21] Fafchamps, Marcel, Christopher Udry and Katherine Czukas. 1998. "Drought and saving in West Africa: are livestock a buffer stock?" *Journal of Development Economics*, 55: 273-305.

[22] Field, Erica, and Attila Ambrus. 2008. "Early marriage, age of menarche, and female schooling attainment in Bangladesh." *Journal of Political Economy* 116(5): 881-930.

[23] Foster, Andrew D. and Mark R. Rosenzweig. 1996. "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution," *American Economic Review*, 86(4): 931-953.

[24] Geertz, Hildred. 1961. *The Javanese Family: A Study of Kinship and Socialization.* New York: The Free Press of Glencoe, Inc.

[25] Giné, Xavier, Robert Townsend, and James Vickery. 2008. "Patterns of Rainfall Insurance Participation in Rural India." *World Bank Economic Review,* 22(3): 539–66.

[26] Gertler, Paul and Jonathan Gruber. 2002. ""Insuring Consumption Against Illness." *American Economic Review*, 92(1): 51-70.

[27] Greif, Avner. 2006. "Family Structure, Institutions, and Growth: The Origins and Implications of Western Corporations." *American Economic Review*, 96(2): 308-312.

[28] Goldin, Claudia. 1995. *The U-Shaped Female Labor Force Function in Economic Development and Economic History. Investment in Women's Human Capital and Economic Development.* (T. P. Schultz, Ed.).: 61-90.: University of Chicago Press.

[29] Goode, William J. 1964. *The family.* New Jersey: Prentice-Hall, Inc.

[30] Goode, William J. 1970. *World Revolution and Family Patterns.* New York: Free Press.

[31] Goody, Jack. 1983. *The development of the family and marriage in Europe.* Cambridge: Cambridge University Press.

[32] Hasil sensus sampel (Result of sample census). 1986. Biro Pusat Statistik: Jakarta.

[33] Hazell, Peter, Jamie Anderson, Niels Balzer, Andreas Hastrup Clemmensen, Ulrich Hess and Francesco Rispoli. 2010. "Potential for Scale and Sustainability in Weather Index Insurance for Agriculture and Rural Livelihoods." *International Fund for Agricultural Development and World Food Programme.* Rome, March.

[34] Hess, Gregory D. 2004. "Marriage and Consumption Insurance: What's Love Got to Do with It?," *Journal of Political Economy*, 112(2): 290-318.

[35] Hill, Hal. 2000. *The Indonesian Economy.* Cambridge: Cambridge University Press.

[36] Jacoby, Hanan G. and Emmanuel Skoufias. 1997. "Risk, Financial Markets, and Human Capital in a Developing Country." *The Review of Economic Studies*, 64(3): 311-335.

[37] Jayachandran, Seema. 2006. "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries." *Journal of Political Economy*, 114(3): 538-575.

[38] Jensen, Robert (2002). "Fertility Preferences and Female Disadvantage: Equal Treatment, Unequal Outcomes?" mimeo, John F. Kennedy School of Government, Harvard University, Cambridge.

[39] Jensen, Robert. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *Quarterly Journal of Economics*. 127(2): 753-792.

[40] Jones, Gavin. 1981. "Malay Marriage and Divorce in Peninsular Malaysia: Three Decades of Change." *Population and Development Review*, 7(2): 255-278.

[41] Jones, Gavin, 1997. "Modernization and Divorce: Contrasting Trends in Islamic Southeast Asia and the West." *Population and Development Review*, 23(1): 95-114.

[42] Jones, Gavin. 2010. "Changing Marriage Patterns in Asia". Asia Research Institute Working Paper No. 131.

[43] Klonner, Stefan. 2003. "Rotating Savings and Credit Associations When Participants Are Risk Averse," *International Economic Review*, 44(3): 979-1005.

[44] Kompilasi data output & dan input usaha tani padi intensifikasi per kabupaten di Jawa-Madura. 1976/77. Biro Pusat Statistik: Jakarta.

[45] Kompilasi data luas panen, hasil per hektar, dan produksi padi terperinci menurut jenis intensifikasi per propinsi di Indonesia. 1980. Biro Pusat Statistik: Jakarta.

[46] Kompilasi data luas panen, hasil per hektar dan produksi padi terperinci menurut jenis intensifikasi MT 1980 1981 dan MT 1981 per kabupaten di Jawa dan Madura. 1982. Biro Pusat Statistik: Jakarta.

[47] Lingon, Ethan, Jonathan P. Thomas and Tim Worral. 2002. "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies," *Review of Economic Studies*, 69(1): 209-44.

[48] Luas panen, rata-rata hasil, dan produksi tanaman padi dan palawija, 1980 per-kabupaten. 1984. Direktorat Jenderal Pertanian Tanaman Pangan Direktorat Bina Program Tanaman Pangan Sub Dit. Data dan Statistik: Jakarta.

[49] Luas tanah menurut penggunaannya di luar Jawa. 1980. Biro Pusat Statistik: Jakarta.

[50] Luas tanah menurut penggunaannya di Jawa & Madura. 1980. Biro Pusat Statistik: Jakarta.

[51] Lucas, Robert E. B. and Oded Stark. 1985. "Motivations to Remit: Evidence from Botswana," *Journal of Political Economy*, 93(5): 901-918.

[52] Mace, Barbara. 1991. "Full Insurance in the Presence of Aggregate Uncertainty," *Joumal of Political Economy*, 99: 928-956.

[53] Maccini, Sharon, and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review*, 99(3): 1006–26.

[54] Manning, Chris. 1988. *Green revolution, employment, and economic change in rural Java : a reassessment of trends under the New Order*. Institute of Southeast Asian Studies: Singapore.

[55] Mazzocco, Maurizio, and Shiv Saini. 2012. "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences." *American Economic Review*, 102(1): 428-68.

[56] Mobarak, Ahmed Mushfiq and Mark Rosenzweig. 2012. ”Selling Formal Insurance to the Informally Insured,” Working Papers 1007, Economic Growth Center, Yale University.

[57] Morduch, Jonathan. 1994. “Poverty and Vulnerability,” *American Economic Review P&P*, 84(2): 221-225.

[58] Morduch, Jonathan. 1995. “AssociationIncome Smoothing and Consumption Smoothing,” *Journal of Economic Perspectives*, 9(3): 103-114.

[59] Morten, Melanie. 2013. ““Temporary Migration and Endogenous Risk Sharing in Village India,” mimeo, Yale University.

[60] Munshi, Kaivan and Mark Rosenzweig. 2006. “Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy.” *American Economic Review* 96(4):1225-1252.

[61] Munshi, Kaivan and Mark Rosenzweig. 2009. ”Why is Mobility in India so Low? Social Insurance, Inequality, and Growth”. NBER Working Papers 14850, National Bureau of Economic Research, Inc.

[62] Palmer, Ingrid. 1976. *New rice in Indonesia*. United Nations Research Institute for Social Development.

[63] Paxson, Christina H. 1992. “Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand.” *American Economic Review*, 82(1): 15-33.

[64] Pitt, Mark, Mark Rosenzweig and Donna Gibbons. 1993. “The Determinants and Consequences of Placement of Government Programs in Indonesia,” *World Bank Economic Review*, 7: 319-348.

[65] Ravallion, Martin and Shubham Chaudhuri.1997. “Risk and Insurance in Village India: Comment”. *Econometrica*, 65(1): 171-184.

[66] Resosudarmo, Budy and Satoshi Yamazaki. 2011. ”Training and Visit (T&V) Extension vs. Farmer field School: The Indonesian Experience,” Departmental Working Papers 2011-01, The Australian National University.

[67] Rose, Elaina. 1999. “Consumption Smoothing and Excess Female Mortality in Rural India.” *Review of Economics and Statistics*, 81(1): 41–49.

[68] Rosenzweig, Mark R. and Hans P. Binswanger. 1993. “Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments.” *The Economic Journal*, 103(416): 56-78.

[69] Rosenzweig, Mark R & Stark, Oded, 1989. ”Consumption Smoothing, Migration, and Marriage: Evidence from Rural India.” *Journal of Political Economy*, 97(4): 905-26.

[70] Rosenzweig, Mark R. and Kenneth I. Wolpin. 1993. “Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India.” *Journal of Political Economy*, 101(2): 223-244.

[71] Rubio, Gabriela. 2013. “The Love Revolution: Decline in Arranged Marriages in Asia, the Middle East and Sub-Saharan Africa” mimeo, University of California, Los Angeles.

[72] Sensus pertanian Republik Indonesia, 1963: penerbitan terachir. 1963. Biro Pusat Statistik: Jakarta.

[73] Shore, Stephen H. 2010. “For Better, For Worse: Intra-Household Risk-Sharing over the Business Cycle,” *Review of Economics and Statistics*, 92(3): 536–548.

[74] Stiglitz, Joseph E. 1974. “Incentives and Risk Sharing in Sharecropping.” *The Review of Economic Studies*, 41(2) : 219-255.

[75] Strauss, John and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions," in Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics.* Amsterdam: Elsevier Science.

[76] Ter Haar, B. 1948. *Adat Law in Indonesia.* New York: Institute of Pacific Relations.

[77] Townsend, Robert. 1994, "Risk and Insurance in Village India." *Econometrica*, 62: 539-591.

[78] Udry, Christopher. 1994. "Risk and Insurance in Rural Credit Market_An Empirical Investigation in Northern Nigeria." *The Review of Economic Studies*, 61(3) : 495-526.

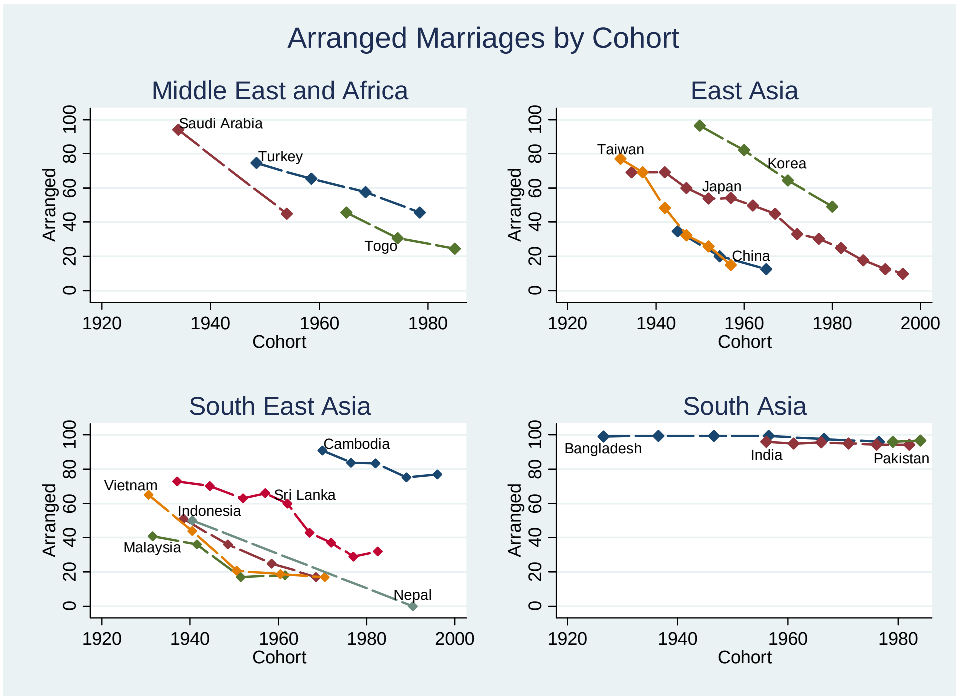
[79] Vreede-de Stuers, Cora. 1960. *The Indonesian Woman: Struggles and Achievements.* The Hague: Mouton & Co., Publishers.

[80] Vogl, Tom S. 2013. "Marriage Institutions and Sibling Competition: Evidence from South Asia." *Quarterly Journal of Economics*, 128 (3): 1017-1072.

[81] Zaman, Muhammad. 2008. "Socio-cultural Security, Emotions and Exchange Marriages in an Agrarian Community." SAGE Publications 28(3): 285–298.

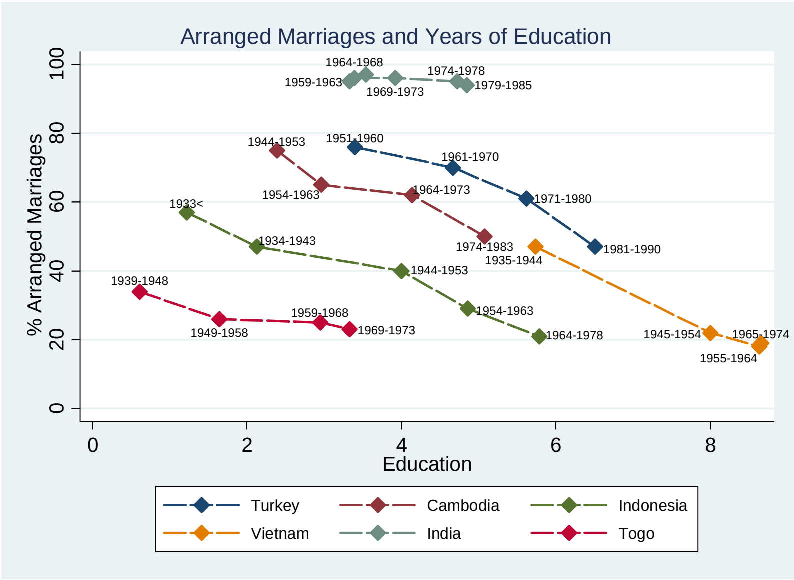
7 Graphs and Tables

Figure 1: Arranged Marriages by Cohort and Region



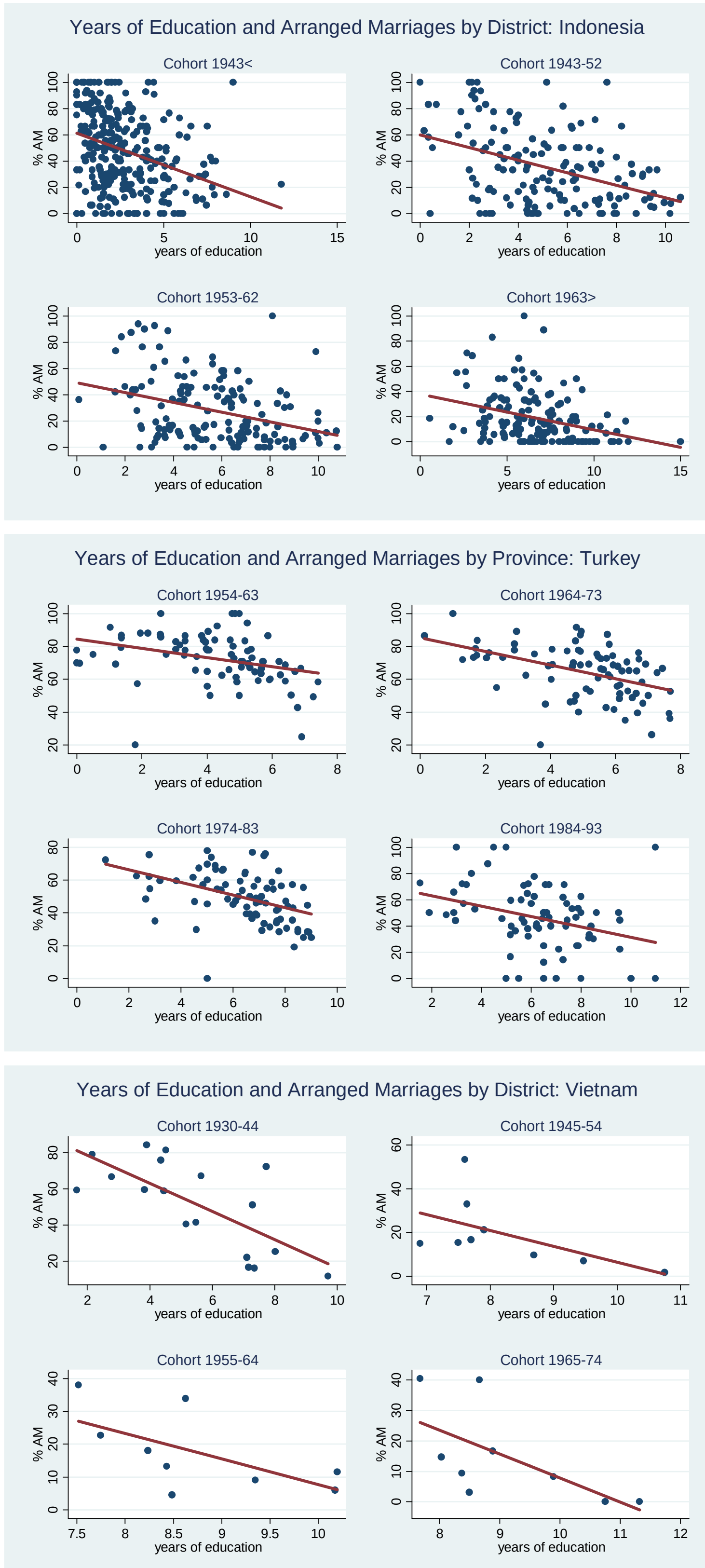
The data sources used to generate figures 1 to 6 are described in detail in Rubio (2013). For the following countries, I have used aggregate information collected from survey reports and other research papers: Japanese National Fertility Survey (Japan), Korean National Fertility Survey (Korea), Chitwan Valley Family Study (Nepal), Chengtu City and Urumchi city Surveys (China), Shefar 'Am Arab community (Israel), Coastal Sri Lanka Survey (Sri Lanka), Southern Ghana Survey (Ghana), Asian Marriage Survey (Thailand), City of Damman Survey (Saudi Arabia), Taiwan Provincial Institute of Family Planning (Taiwan), Malaysian Marriage Survey (Malaysia). For the rest of the countries I have used micro-data from: Cambodian (2000 and 2005), Togolese (1988) and Turkish (1998, 2003 and 2008) Demographic and Health Survey, Vietnam Longitudinal Survey (1995-1998), Indonesia Family Life Survey (1993, 1997, 2000, 2007), India Human Development Survey (2005), and Matlab (Bangladesh) Health and Socio-Economic Survey (1996).

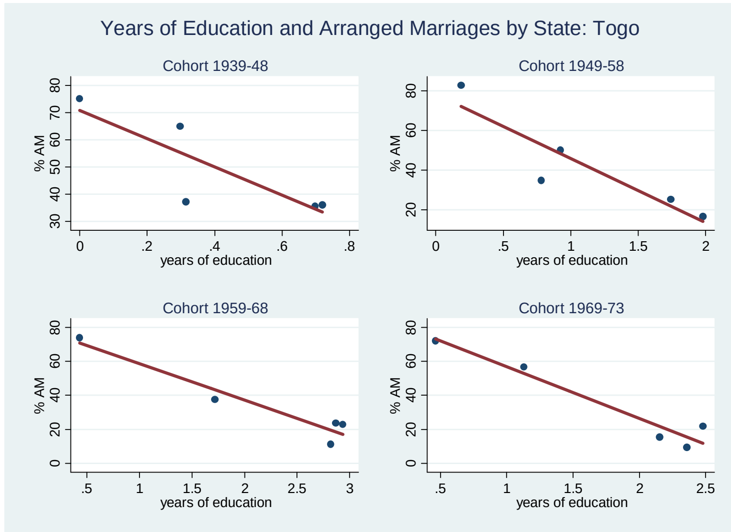
Figure 2: Arranged Marriages and Education by Cohort

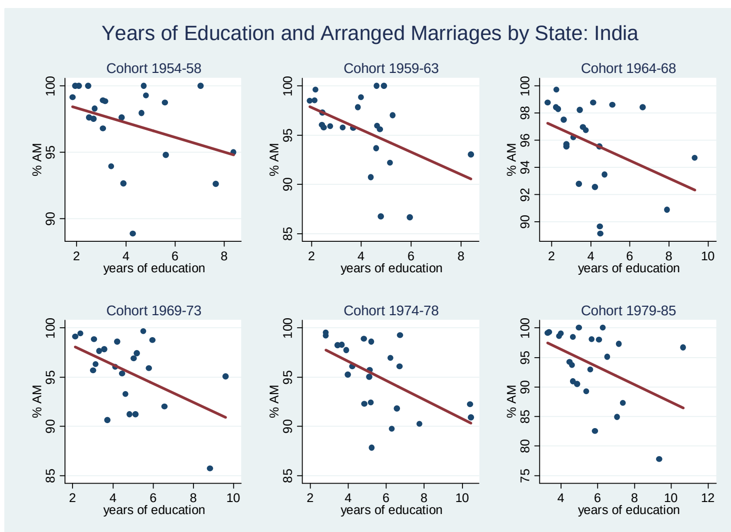


Each line of the graph refers to a different country and each point represents a different cohort for women. It correlates the average years of schooling of the cohort with the average percentage of women who ever had an arranged marriage within the same cohort.

Figure 3: Arranged Marriages and Years of Schooling within Cohort: Geographic Variation

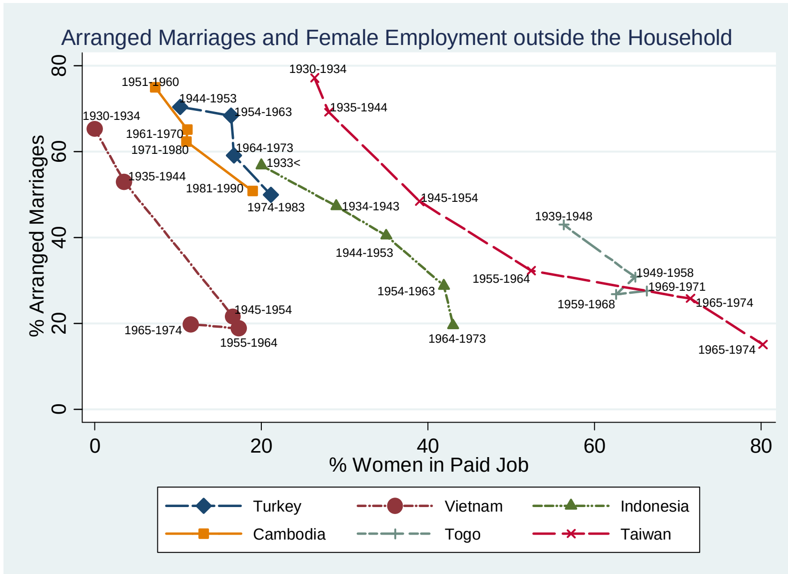






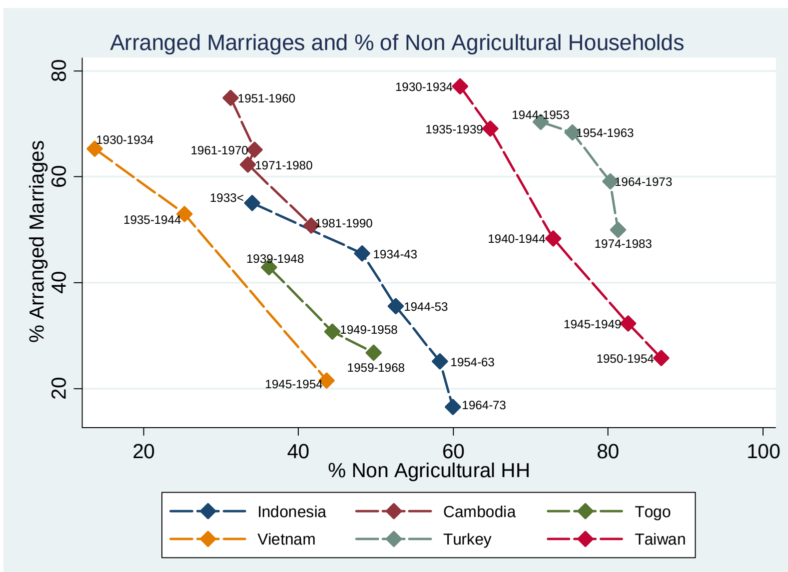
The graph for Indonesia uses data from 148 districts belonging to the provinces of Sumatera Utara, Sumatera Barat, Suamtera Selatan, Lampung, Dki Jakarta, Jawa Barat, Jawa Tengah, Di Yogyakarta, Jawa Timur, Bali, Nusa Tenggara Barat, Kalimantan Selatan and Sulawesi Selatan. For Turkey, the geographic variation comes from the provinces Adana, Adiyaman, Afyon, Agri, Amasya, Anakara, Antalya, Artvin, Aydin, Balikesir, Bilecik, Bingöl, Bitlis, Bolu, Burdur, Bursa, Çanakkale, Çankiri, Çorum, Denizli, Diyarbakir, Edirne, Elazig, Erzingan, Erzurum, Eskisekir, Gaziantep, Giresun, Gümüshane, Hakkari, Hatay, Isparta, Içel, Istanbul, Izmir, Kars, Kastamonu, Kayseri, Kirklareli, Kirsehir, Kocaeli, Konya, Kütahya, Malatya, Manisa. K. Maras, Mardin, Mugla, Mus, Nevsehir, Nigde, Ordu, Rize, Sakarya, Samsun, Siirt, Sinop, Sivas, Tekirdag, Tokat, Trabzon, Tunceli, Sanliurfa, Usak, Van, Yozgat, Zonguldak, Aksaray, Bayburt, Karaman, Kirikkale, Sirnak, Bartin, Ardahan, Igdir, Yalova, Karabük, Kilis, Osmaniye, and Düzce. For Vietnam, the variation used is across districts: Binh Luc, Hai Hau, Nam Dinh, Nam Ninh, Ninh Binh, Phu Ly, Thanh Liem, Xuan Thuy, and Yen Khang. For Togo, I used the following regions: Maritime, Des Plateaux, Centrale, De la Kara, and Des Savanes. And finally for India, I focused on differences across states: Jammu & Kashmir, Himachal Pradesh, Punjab, Uttaranchal, Haryana, Delhi, Rajasthan, Uttar Pradesh, Bihar, Assam, West Bengal, Jharkhand, Orissa, Chhatishgarh, Madhya Pradesh, Gujarat, Maharashtra, Andhra Pradesh, Karnataka, Goa, Kerala, and Tamil Nadu.

Figure 4: Arranged Marriages and Female Employment



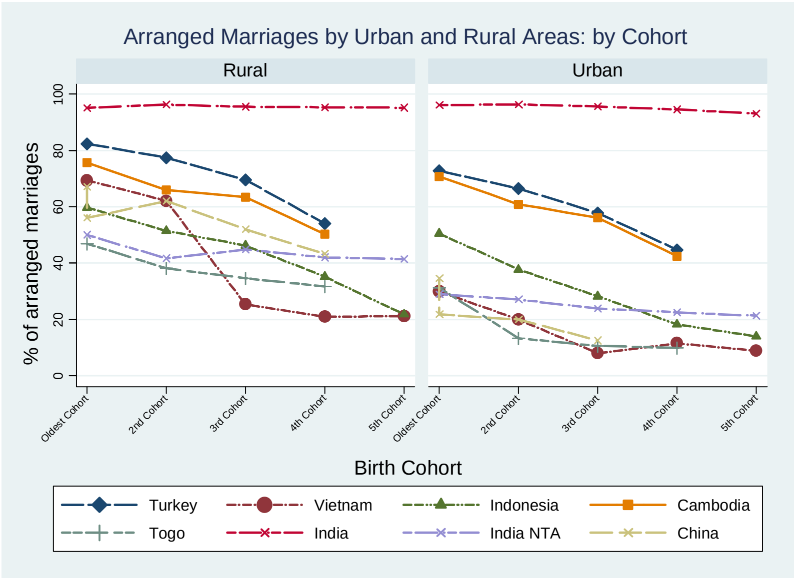
The definition of employment outside the household varies for each country depending on the information available. For Cambodia, Indonesia, and Turkey, women report three types of work status: employee, self-employed and family worker. This figure plots the percentage of employees by cohort. For Togo, there is no information on the work status of women; the variable used instead is the percentage of women "owning" their wages conditional on being in the labor force (women who can freely spend their wages). For Vietnam, the variable used corresponds to women working outside the household for non-relatives conditional on being in the labor force. For Taiwan, I use aggregate information on work status reported by cohort.

Figure 5: Arranged Marriages and Percentage of Non-Agricultural Households



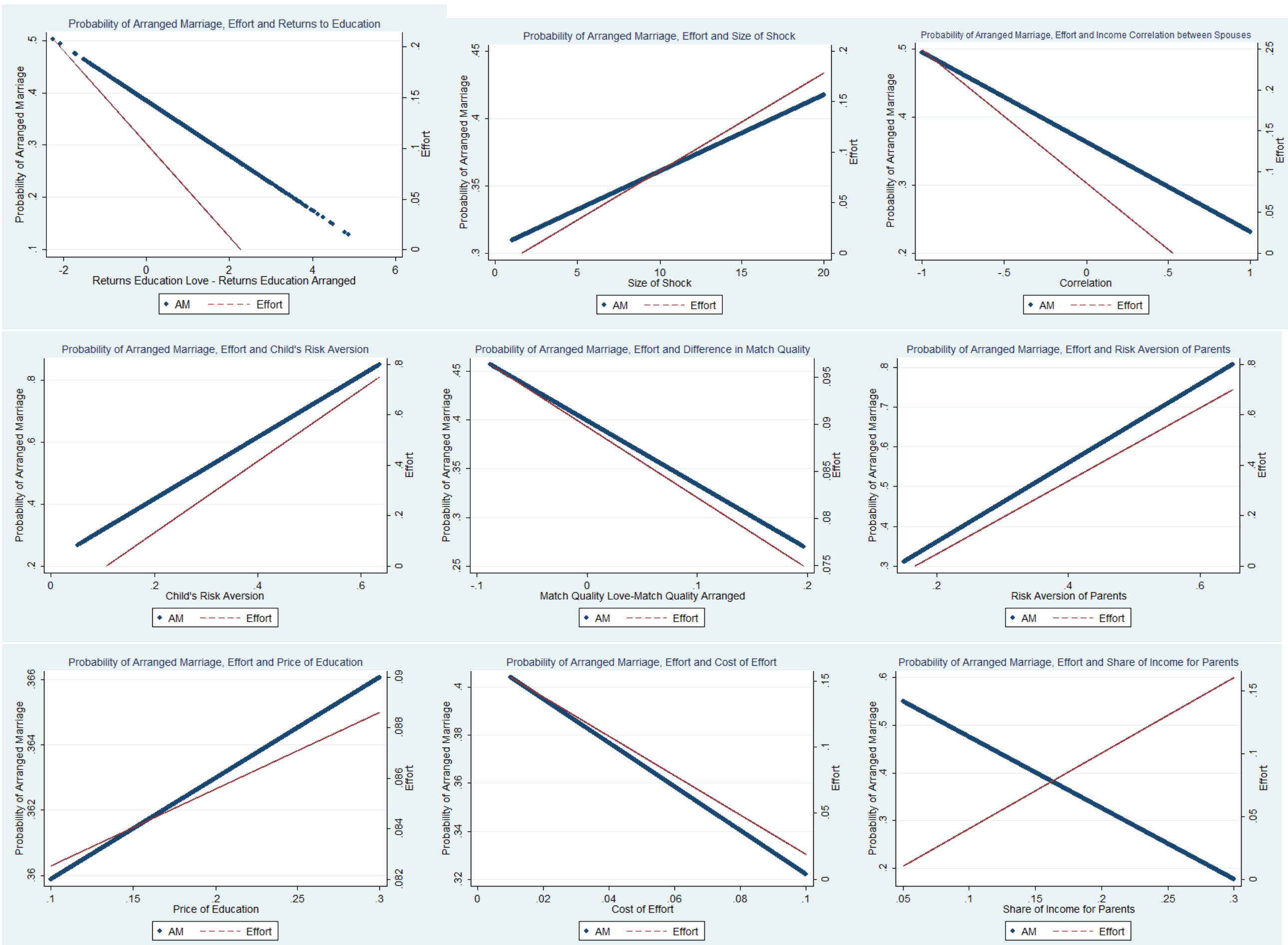
Non-agricultural refers to households whose main income source is not agriculture.

Figure 6: Arranged Marriages by Cohort and Residence



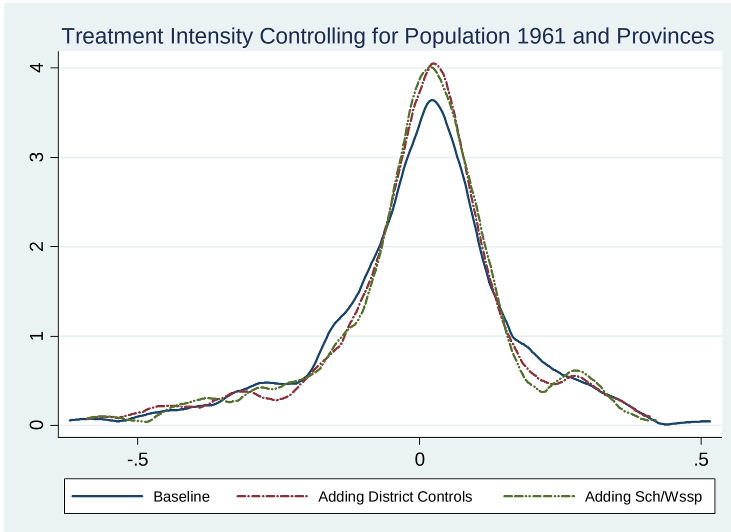
For Turkey, the cohorts are 1944-53, 1954-63, 1964-73, and 1974-83; for Vietnam, 1930-34, 1935-44, 1945-54, 1955-64 and 1965-74; for Indonesia,1933<, 1934-43, 1944-53, 1954-63 and 1964-78; for Cambodia, 1951-60, 1961-70, 1971-80 and 1981-90; for Togo, 1939-48, 1949-58, 1959-68 and 1969-73; for India, 1959-63, 1964-68, 1969-73, 1974-78 and 1979-85.

Figure 7: Numerical Comparative Statics: One child



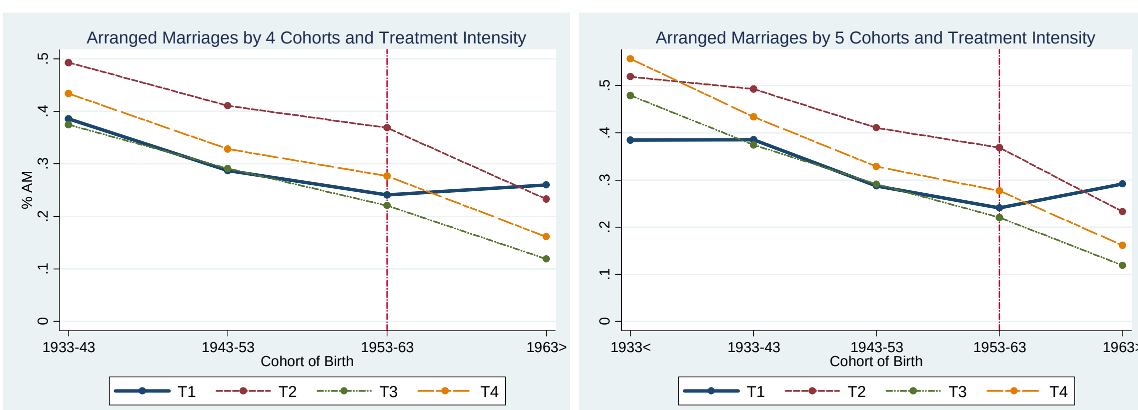
The parameters used for the simulations are $p \in [0.1, 0.3]$, $\beta \in [0.8, 1]$, $\varphi \in [0.05, 0.3]$, $d \in [0.15, 0.65]$, $d_k \in [0.01, 0.55]$, $e_{high} \in [0.01, 0.25]$, $e_{low} = 0$, $\sigma_\delta^2 \in [0.01, 1.2]$, $\varrho_{ks}\,(I(e=1),\, I(L=1)) \in [-1, 1]$, $\alpha_L \in [0.01, 0.2]$, and $\alpha_A \in [0.001, 0.1]$; x_L and x_A are bounded such that education is increasing in the returns.

Figure 8: Plotting Residuals of Treatment Intensity



The baseline figure plots the density of the residuals of the treatment variable after controlling for the total population by district in 1961 and province fixed effects. It then adds controls for district characteristics of 1971. And finally, it controls for the intensity of the school construction, and the water and sanitation supply program. The density does not change as I add information at the district level suggesting that the intensity of the treatment in the early 1980s is not correlated with district characteristics in 1970s. The results are also shown in appendix table 20.

Figure 9: Arranged Marriages by Cohort of Birth and Treatment Intensity



I divided the treatment variable into the four quartiles of the distribution of sawah land covered by any intensitifcation program. The figure plots the mean AM by cohort of birth and treatment intensity, where treatment one corresponds to the lowest treatment (the districts with the lowest percentage of sawah land covered by the GR). The figure shows that cohorts 1933-42, 1943-52 and 1953-62 follow the same trends in the four treatment intensity areas. It also shows the that identication is obtained from the last cohort by comparing areas more intensily treated against areas with low treatment intensity (after differencing out the common pre-trends).

Table 1: Demographics, Mean and Median Per Capita Expenditure and Income

Demographics by type of marriage						
1993	Type of Marriage		Love	Arranged	Mixed=1 Love, 1 Arranged	
	Observations		1438	410	313	
	Yrs School Husband		6.30	4.14	5.01	
	Yrs School Wife		5.12	3.48	4.16	
	1993		1997		2000	
	Love	Arranged	Love	Arranged	Love	Arranged
Age	39.60	47.00	43.43	51.27	46.40	54.04
AE size	3.94	4.22	4.02	4.23	4.02	4.24
Urban	37.9%	27.5%	38.6%	27.5%	37.9%	29.0%
Farmer	46.9%	63.3%	42.6%	55.2%	52.7%	65.5%
Business	35.0%	36.3%	37.5%	34.3%	49.3%	44.3%
Mean Monthly per capita Expenditure by type of marriage						
	1993		1997		2000	
	Love	Arranged	Love	Arranged	Love	Arranged
Food	31.3	24.5	38.9	32.1	41.0	33.0
Exp1	7.5	4.5	10.5	6.2	9.7	6.1
Exp2	4.0	2.5	4.1	3.1	4.4	4.8
Total	43.8	31.7	52.5	40.4	53.4	41.3
Median Monthly per capita Expenditure by type of marriage						
	1993		1997		2000	
	Love	Arranged	Love	Arranged	Love	Arranged
Food	23.8	18.4	30.1	23.6	31.7	27.2
Exp1	3.7	2.3	4.3	3.2	4.7	3.4
Exp2	1.9	1.3	2.4	1.7	2.3	1.7
Total	32.6	24.2	38.9	29.0	41.0	33.5
Mean Monthly per capita Income by type of marriage						
	1993		1997		2000	
	Love	Arranged	Love	Arranged	Love	Arranged
Wages + Profits	89.8	118.2	89.2	43.9	80.1	59.3
All income	170.0	186.9	109.4	49.8	85.4	64.7
Median Monthly per capita Income by type of marriage						
	1993		1997		2000	
	Love	Arranged	Love	Arranged	Love	Arranged
Wages + Profits	31.7	20.1	40.2	25.7	45.8	37.4
All income	32.9	21.1	42.0	29.2	48.6	40.3

Mean Monthly Transfers by type of marriage						
Transfers out HH	53.0	26.8	18.0	11.9	34.7	70.9
Transfers in HH	24.4	12.2	11.2	11.9	12.3	12.6
% Transfers out HH*	86%	76%	84%	76%	90%	89%
% Transfers in HH*	67%	62%	70%	75%	82%	86%
As % of Total Income Per Capita						
Transfers out HH	31.2%	14.3%	16.5%	23.9%	40.6%	109.6%
Transfers in HH	14.3%	6.5%	10.2%	23.8%	14.4%	19.5%
(Means of) Formal Insurance and Other NLI						
	1993		1997		2000	
	Love	Arranged	Love	Arranged	Love	Arranged
% Insurance**	21%	10%	18%	6%	18%	7%
% Savings ***	34%	25%	31%	22%	43%	33%
Savings Amount	1362.0	176.7	373.2	256.8	574.5	622.5
Debt	372.7	202.4	445.7	171.7	254.8	127.8
Pensions	296.0	307.4	60.8	23.6	19.7	18.3
Other NLI	180.3	173.7	34.3	27.3	18.5	15.2
As % of Household Monthly Total Income						
Savings Amount	3.97	1.93	0.68	0.41	1.26	1.28
Debt	2.75	1.68	1.48	1.10	0.64	0.51
Pensions	0.04	0.03	0.05	0.06	0.04	0.04
Other NLI	0.41	1.10	0.12	0.13	0.05	0.06
Transfers to and from parents and parents-in-law as % of Household Monthly Total Income						
To Parents	13.3%	2.0%	3.2%	2.9%	17.5%	20.7%
From Parents	4.6%	3.8%	1.5%	1.2%	0.9%	0.5%

Exp 1= utilities + personal toiletries + small household items.

Exp 2 = recreation and entertaiment + transportation + clothing + taxes.

*% of Households where the head or his wife report receiving/giving a transfer; **% of Households where at least one member reports having any type of formal insurance; ***% of Households reporting having savings.

I use the total amount of wages reported within a household (from all household members); and the total profit for agricultural and non-agricultural business reported in the household roster and in the individual questionnaire. Non-labor income includes pensions, rents from assets and other bonuses. Expenditure on non-durables includes food expenses, utilities, personal toiletries, small non-durable household items, recreation, entertainment, transportation, clothing, taxes and rent. Transfers out (in) refers to the monthly amount of transfers given (received) to (from) parents and siblings by the household head or his wife. Savings amount refers to the total self-reported amount of savings, certificate of deposit, stocks and receivables. Debt is the total outstanding debt value at the household level. Other non-labor income (not included as part of all income) includes scholarships, insurance claims, arisan, and gifts reported in the household roster and in the individual questionnaire. % Insurance refers to the percentage of households where at least one member reports having formal insurance (provided by employer or other). % Savings is the percentage of households that report having savings.

Table 2: Full Insurance Tests by Type of Marriage: First Differences and Fixed Effects

	Inverse hyperbolic sine transformation				**Neglog transformation**			
	First Differences		Fixed Effects		First Differences		Fixed Effects	
	Arranged	Love	Arranged	Love	Arranged	Love	Arranged	Love
Log Agg Cons	0.469***	0.342***	0.484***	0.461***	0.483***	0.357***	0.497***	0.474***
	(0.0856)	(0.0512)	(0.0731)	(0.0437)	(0.0865)	(0.0516)	(0.0739)	(0.0439)
Log Income	**-0.00298**	**0.00913****	**0.00120**	**0.00852****	**-0.00303**	**0.00977****	**0.00151**	**0.00911****
	(0.00550)	(0.00392)	(0.00547)	(0.00341)	(0.00584)	(0.00412)	(0.00581)	(0.00358)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	820	2,876	1,230	4,314	820	2,876	1,230	4,314
R-squared	0.051	0.087	0.154	0.171	0.052	0.088	0.155	0.172
Number of hhid			410	1,438			410	1,438
Joint Test (Log Agg Cons = 1) (Log Income = 0)								
F-statistic	19.49	85.04	24.93	79.48	18.08	80.02	23.26	75.12
P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Some households report losses in profits or zero income in some components. I use two alternative transformations used in the literature instead of directly using a logarithmic transformation: (i) Inverse sine transformation $asinh(y) = log[y_i + (y_i^2 + 1)^{1/2}]$; and (ii) the neglog transformation $neglog(y) = \begin{cases} log(y+1) & if\ y \geq 0 \\ -log(1-y) & if\ y < 0 \end{cases}$. The results found under either transformation are very similar.

Additional controls are age of the household head, number of adult equivalent members, a dummy variable for residence and a dummy variable for agricultural activities. Standard errors clustered at the household level.

Table 3: Effect of the Green Revolution on Returns to Education

Panel A. Effect on Wages using SUSENAS and Census

Dep. Var: ln wage						
	BIMAS/INMAS		**Reduced Form**		**IV**	**IV**
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling*BIMAS/INMAS	**0.0369*** **	**0.0306*** **			**0.0671*** **	**0.0551*** **
	(0.00545)	(0.00539)			(0.0124)	(0.0132)
Years of schooling*Sawah Area			**0.0605*** **	**0.0495***		
			(0.0231)	(0.0257)		
Years of schooling*Farms per Ha			**0.0409*** **	**0.0349*** **		
			(0.00789)	(0.00838)		
Years of schooling*% Sawah Area* Farms/Ha			**-0.0572*** **	**-0.0512** **		
			(0.0202)	(0.0227)		
Years of schooling	0.0544***	0.0596***	0.0407***	0.0497***	0.0289***	0.0379***
	(0.00418)	(0.00413)	(0.00763)	(0.00784)	(0.0104)	(0.0114)
Constant	7.918***	8.111***	8.265***	8.153***	7.954***	6.023***
	(0.0551)	(0.155)	(0.0659)	(0.275)	(0.0666)	(0.418)
District FE	yes	yes	yes	yes	yes	yes
Interaction with Islands Dummy	no	yes	no	yes	no	yes
Observations	131,793	131,793	131,793	131,793	131,793	131,793
R-squared	0.724	0.726	0.725	0.726	0.724	0.726
F-statistic of excluded instruments					25.43	16.13

Panel B. Effect on Wages using IFLS

	BIMAS/INMAS		**Reduced Form**		**IV**	**IV**
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling*BIMAS/INMAS	**0.0557***	**0.0527** **			**0.211** **	**0.200** **
	(0.0260)	(0.0238)			(0.105)	(0.101)
Years of schooling*Sawah Area			**0.0127**	**0.0570**		
			(0.0701)	(0.0639)		
Years of schooling*Farms per Ha			**0.0145**	**0.0368**		
			(0.0320)	(0.0320)		
Years of schooling*% Sawah Area* Farms/Ha			**-0.0110**	**-0.0437**		
			(0.0621)	(0.0615)		
Years of schooling	0.0806**	0.0744***	0.109***	0.0656**	-0.0575	-0.0576
	(0.0266)	(0.0205)	(0.0348)	(0.0325)	(0.0943)	(0.0917)
Constant	2.743***	3.063***	2.542***	3.622***	2.709***	1.170***
	(0.228)	(0.266)	(0.361)	(0.379)	(0.950)	(0.259)
Province FE	yes	no	yes	no	yes	no
District FE	no	yes	no	yes	no	yes
Observations	2,394	2,394	2,730	2,730	2,298	2,298
R-squared	0.371	0.428	0.414	0.476	0.352	0.415
F-statistic of excluded instruments					3.54	3.74

Standard errors clustered at the district level

Panel A. Years of education were imputed for 1980, 1981 and 1982 using self-reported information on highest level of education completed. 1976 and 1995 report the number of years of education. Panel B. The variable on years of education was constructed using information on highest level of education and grade completed (see data appendix for other details on the samples). Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83; Sawah Area = % of agricultural land available for production of wet paddy (rice) in 1963; Farms/Ha = Total number of farms in 1963/ Total Agricultural Land (Ha) in 1963.

Table 4: Effect of the Green Revolution on Mean and Variance of Income

Panel A. Effect on per capita Income using SUSENAS and Census

MLE effect on log pc income mean and variance: Agricultural Households

	Effect of BIMAS/INMAS				**Reduced Form of Sawah Land and Farms/ha**			
	Income	Variance	Income	Variance	Income	Variance	Income	Variance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Area BIMAS/INMAS	**0.0960**	**0.0420**	**0.117**	**-0.118**				
	(0.0826)	(0.0880)	(0.106)	(0.189)				
% Sawah Area					**-0.0717**	**-0.481**	**0.301**	**-0.985**
					(0.241)	(0.451)	(0.323)	(0.646)
Number farms per Ha					**-0.328*****	**-0.249**	**-0.224**	**-0.402**
					(0.105)	(0.177)	(0.137)	(0.253)
% Sawah Area* Farms/Ha					**0.557*****	**0.406**	**0.313**	**0.736**
					(0.204)	(0.351)	(0.270)	(0.486)
Rural	-0.200***	-0.0137	-0.200***	-0.0147	-0.196***	-0.0133	-0.200***	-0.0138
	(0.0218)	(0.0185)	(0.0219)	(0.0198)	(0.0214)	(0.0188)	(0.0212)	(0.0196)
Constant	10.28***	0.783***	10.26***	0.918***	10.44***	1.062***	10.30***	1.286***
	(0.0631)	(0.0831)	(0.0829)	(0.181)	(0.122)	(0.233)	(0.149)	(0.348)
Province FE	yes	yes	yes	yes	yes	yes	yes	yes
Interaction with Islands Dummy	no	no	yes	yes	no	no	yes	yes
Observations	48,079	48,079	48,079	48,079	48,079	48,079	48,079	48,079
Wald chi2								

Panel B. Effect on per capita Income using IFLS

	Income	**Variance**	**Income**	**Variance**	**Income**	**Variance**	**Income**	**Variance**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Area BIMAS/INMAS	**1.245****	**-1.390**	**1.423*****	**-0.545**				
	(0.525)	(1.161)	(0.513)	(0.680)				
Age of household head	-0.0128***	-0.00574	-0.0132***	-0.00438	-0.0116***	-0.000290	-0.0110***	0.00333
	(0.00316)	(0.00850)	(0.00303)	(0.00831)	(0.00282)	(0.00893)	(0.00259)	(0.00696)
Number of household members	-0.0488	0.0536	-0.0334	0.0327	-0.0468	0.0278	-0.0456	-0.0218
	(0.0420)	(0.0596)	(0.0413)	(0.0634)	(0.0421)	(0.0763)	(0.0334)	(0.0493)
% Sawah Area					**2.534**	**-1.971**	**3.561****	**-5.862***
					(1.834)	(4.862)	(1.434)	(3.321)
Number farms per Ha					**0.573**	**-0.660**	**0.695**	**-0.927**
					(0.640)	(0.995)	(0.703)	(1.719)
% Sawah Area* Farms/Ha					**-0.676**	**1.367**	**-1.179**	**3.318**
					(1.320)	(3.045)	(1.240)	(3.341)
Constant	10.06***	3.087***	9.876***	2.345***	10.16***	2.317**	9.956***	2.961*
	(0.602)	(1.128)	(0.537)	(0.759)	(0.775)	(0.942)	(0.792)	(1.608)
Province FE	yes	yes	yes	yes	yes	yes	yes	yes
Interaction with Islands Dummy	no	no	yes	yes	no	no	yes	yes
Observations	2,505	2,505	2,505	2,505	2,611	2,611	2,611	2,611
Wald chi2	111.81		126.34		215.12		210.29	

Standard errors clustered at district level

See data appendix for details on the samples. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83; Sawah Area = % of agricultural land available for production of wet paddy (rice) in 1963; Farms/Ha = Total number of farms in 1963/ Total Agricultural Land (Ha) in 1963.

Table 5: Effect of the Green Revolution on Arranged Marriages

Sample and Results using the introduction of BIMAS/INMAS program

Panel A. Mean Characteristics of Sample merged with BIMAS/INMAS program

Cohort	Obs	Arranged	Female	Education	Rural
1933<	1,185	51.5%	47.5%	2.10	64.8%
1933-1942	1,679	41.9%	54.1%	3.14	65.5%
1943-1952	1,803	32.5%	51.0%	4.85	60.9%
1953-1962	2,488	26.8%	53.3%	5.23	59.8%
1963>	1,913	16.6%	67.7%	5.97	64.1%
Total	9,068	31.8%	55.3%	4.52	62.6%

Panel B. OLS estimates

	Dep. Var: Arranged Marriage							
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
cohort 33-42					-0.0526	-0.358***	-0.421***	-0.304***
					(0.0792)	(0.109)	(0.117)	(0.101)
cohort 43-52	-0.0414	-0.0720	-0.104	-0.0847	-0.102	-0.416***	-0.513***	-0.380***
	(0.0565)	(0.113)	(0.116)	(0.111)	(0.0637)	(0.148)	(0.160)	(0.127)
cohort 53-62	-0.189***	-0.296*	-0.343**	-0.291**	-0.242***	-0.646***	-0.753***	-0.591***
	(0.0710)	(0.150)	(0.151)	(0.146)	(0.0869)	(0.184)	(0.198)	(0.169)
cohort 63>	-0.0996	-0.178	-0.301	-0.264	-0.156	-0.532***	-0.713***	-0.570***
	(0.0844)	(0.183)	(0.183)	(0.181)	(0.102)	(0.202)	(0.219)	(0.201)
cohort 33-42*Area BIMAS/INMAS					-0.104	-0.148*	-0.113	-0.0756
					(0.0815)	(0.0861)	(0.0889)	(0.0836)
cohort 43-52*Area BIMAS/INMAS	-0.0308	-0.0246	-0.00152	0.00872	-0.121	-0.159*	-0.101	-0.0556
	(0.0679)	(0.0721)	(0.0698)	(0.0694)	(0.0810)	(0.0873)	(0.0837)	(0.0776)
cohort 53-62*Area BIMAS/INMAS	**0.0434**	**0.00189**	**0.0342**	**0.0630**	**-0.0544**	**-0.141**	**-0.0774**	**-0.0129**
	(0.0922)	(0.106)	(0.106)	(0.103)	(0.109)	(0.131)	(0.130)	(0.119)
cohort 63>*Area BIMAS/INMAS	**-0.206****	**-0.222***	**-0.152**	**-0.134**	**-0.295****	**-0.359****	**-0.259***	**-0.208**
	(0.103)	(0.130)	(0.125)	(0.122)	(0.127)	(0.152)	(0.146)	(0.137)
Female	0.148***	0.148***	0.149***	0.148***	0.151***	0.152***	0.152***	0.152***
	(0.0125)	(0.0126)	(0.0126)	(0.0126)	(0.0125)	(0.0125)	(0.0125)	(0.0125)
Constant	-0.0459	-0.304**	-0.538**	-0.0395	0.00790	-0.361**	-0.422	-0.0486
	(0.0839)	(0.127)	(0.220)	(0.163)	(0.104)	(0.155)	(0.275)	(0.190)
District FE	yes	yes	yes	yes	yes	yes	yes	yes
District Controls	no	yes	yes	yes	no	yes	yes	yes
School Const. Program	no	no	yes	yes	no	no	yes	yes
Water & Sanitation Supply	no	no	no	yes	no	no	no	yes
Observations	7,883	7,883	7,883	7,883	9,068	9,068	9,068	9,068
R-squared	0.264	0.266	0.267	0.268	0.278	0.282	0.283	0.284

Standard errors clustered at the district level

For columns 1 to 4, the omitted cohort is the cohort born 1933-43. For columns 5 to 8, the omitted cohort is the cohort born before 1933. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83.

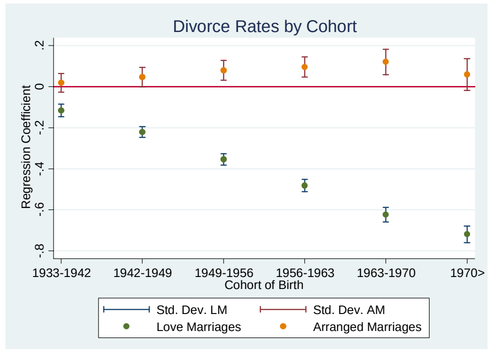
Table 6: Effect of the Green Revolution on Years of Education

Dep. Var: Education	IFLS	IFLS	IFLS	IFLS	IFLS	IFLS
	(1)	(2)	(3)	(4)	(5)	(6)
cohort 33-42				1.383	3.726***	1.016
				(0.880)	(1.295)	(1.047)
cohort 43-52	1.082*	1.217	0.564	2.472***	5.010***	1.566
	(0.591)	(0.910)	(0.821)	(0.905)	(1.584)	(1.102)
cohort 53-62	1.617**	3.721***	2.099***	3.005***	7.412***	3.119**
	(0.664)	(1.166)	(0.762)	(1.064)	(2.073)	(1.228)
cohort 63>	2.233***	3.493***	2.650***	3.585***	7.033***	3.602**
	(0.796)	(1.307)	(0.905)	(1.269)	(2.056)	(1.460)
cohort 33-42*Area BIMAS/INMAS				0.00956	0.169	0.143
				(0.941)	(0.734)	(0.974)
cohort 43-52*Area BIMAS/INMAS	0.940	1.490*	0.680	0.943	1.671*	0.847
	(0.749)	(0.760)	(0.700)	(1.080)	(0.895)	(1.077)
cohort 53-62*Area BIMAS/INMAS	0.691	1.271*	0.162	0.687	1.410*	0.328
	(0.731)	(0.689)	(0.757)	(1.181)	(0.849)	(1.206)
cohort 63>*Area BIMAS/INMAS	**1.465***	**1.782****	**0.937**	**1.485**	**1.916***	**1.149**
	(0.810)	(0.795)	(0.789)	(1.339)	(1.005)	(1.364)
Female	-1.749***	-1.744***	-1.747***	-1.729***	-1.731***	-1.729***
	(0.0875)	(0.0881)	(0.0871)	(0.0789)	(0.0800)	(0.0789)
Constant	8.660***	9.057***	6.815***	7.158***	8.255***	5.718***
	(0.793)	(1.017)	(1.040)	(1.273)	(1.238)	(1.588)
District of birth FE	yes	yes	yes	yes	yes	yes
District Controls	no	yes	yes	no	yes	yes
Int. w/other Programs	no	no	yes	no	no	yes
Observations	7,883	7,883	7,883	9,068	9,068	9,068
R-squared	0.255	0.257	0.257	0.287	0.291	0.289

Standard errors clustered at the district level

For columns 1 to 4, the omitted cohort is the cohort born 1933-43. For columns 5 to 8, the omitted cohort is the cohort born before 1933. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83.

Figure 10: Divorce Trends Indonesia



The coefficients plotted come from the following regression:

$$D_{ipc} = = \beta_0 + \gamma_c + \eta_p + \beta_1 AM_{ipc} + \sum_c (\gamma_c * AM_{ipc})\beta_{2,c} + \sum_p (\eta_p * AM_{ipc})\beta_{3,p} + \beta_4 female_{ipc} + \beta_5 duration_{ipc} + \varepsilon_{ipc}$$

The first graph shows $\beta_{2,c}$, the second graph plots $\beta_{3,p}$. The coefficients are reported in the appendix table 26.

Table 7: Effect of the Green Revolution on the Probability of Divorce for cohorts born before 1952

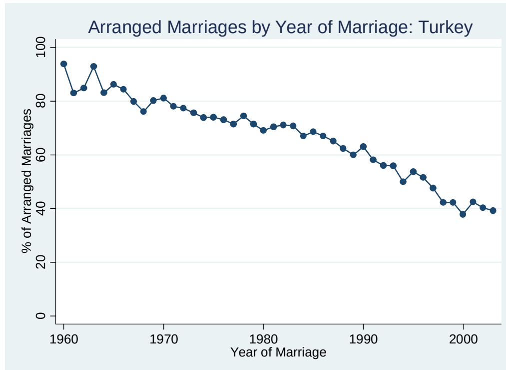
Dep. Var: Divorce

	AM (1)	LM (2)
Cohort 33-42	0.204***	0.136
	(0.0616)	(0.220)
Cohort 43-52	0.280**	-0.0213
	(0.110)	(0.137)
Area BIMAS/INMAS*Married before 1965	**-0.250****	**0.00893**
	(0.108)	(0.233)
Cohort 33-42*Area BIMAS/INMAS*Mbf1965	**0.360*****	**0.123**
	(0.125)	(0.280)
Cohort 43-52*Area BIMAS/INMAS*Mbf1965	**0.421*****	**-0.0635**
	(0.151)	(0.224)
Married before 1965	0.513***	0.300*
	(0.0658)	(0.167)
Cohort 33-42*Married before 1965	-0.339***	-0.207
	(0.0867)	(0.226)
Cohort 43-52*Married before 1965	-0.463***	-0.0279
	(0.110)	(0.162)
Area BIMAS/INMAS	0.681***	0.251
	(0.0957)	(0.206)
Cohort 33-42*Area BIMAS/INMAS	-0.342***	-0.166
	(0.107)	(0.272)
Cohort 43-52*Area BIMAS/INMAS	-0.350**	-0.0309
	(0.149)	(0.209)
Female	0.0130	0.00252
	(0.0140)	(0.0137)
Duration	-0.0181***	-0.0206***
	(0.000630)	(0.000744)
Constant	0.195***	0.296**
	(0.0406)	(0.136)
District FE	yes	yes
Observations	1,800	2,686
R-squared	0.580	0.564

Standard errors clustered at the district level

The omitted cohort is individuals born before 1933. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83. Mbf1965 and Married before 1965 refer to a dummy variable that takes the value of one if the individual was married before 1965.

Figure 11: Percentage of Woman having an Arranged Marriage by Year of Marriage in Turkey



Turkey is another country that has experienced a transition toward love marriages. This graphs plots the (raw) percentage of woman that have an arranged marriage by year of marriage. It shows some differences in the speed of the transition across the years. The decade of the 1960s witnessed a rapid decline (around 15-20%), which slowed down during 1970s and 1980s. The mid-1990s to the late-1990s experienced another phase of accelerated decline which seemed to have also slowed down in the 2000s.

Table 8: Arranged Marriages and Divorce in Turkey

Panel A. Summary Statistics				
	AM		LM	
	Mean	Std Dev	Mean	Std Dev
Age	35.92	7.82	33.31	7.70
Education	4.31	3.20	7.34	4.49
Urban	0.71	0.46	0.80	0.40
Obs	7835		5361	
	Divorce			
Cohort	Mean	Std Dev		
1954-1963	0.018	0.132		
1964-1973	0.024	0.152		
1974-1983	0.023	0.151		
Panel B. Determinants of Divorce				
	Dep. Var: Divorce			
	OLS	**OLS**	**OLS**	**OLS**
	(1)	(2)	(3)	(4)
Age	0.0106***	0.0107***		
	(0.000785)	(0.000813)		
Arranged	**0.0180*** **	**0.0179*** **	**0.00504**	**0.00531**
	(0.00308)	(0.00313)	(0.00486)	(0.00505)
Duration	-0.0107***	-0.0108***	-0.00759***	-0.00760***
	(0.000761)	(0.000785)	(0.000564)	(0.000579)
Cohort 64-73			-0.0814***	-0.0804***
			(0.00833)	(0.00863)
Cohort 74-83			-0.147***	-0.146***
			(0.0130)	(0.0135)
Cohort 64-73*Arranged			**0.0151** **	**0.0150** **
			(0.00633)	(0.00665)
Cohort 74-83*Arranged			**0.0104**	**0.0106**
			(0.00715)	(0.00739)
Constant	-0.218***	-0.228***	0.181***	0.173***
	(0.0194)	(0.0189)	(0.0145)	(0.0138)
District FE	No	Yes	No	Yes
Province FE	Yes	No	Yes	No
Observations	13,196	13,196	13,196	13,196
R-squared	0.133	0.164	0.104	0.136

Standard errors clustered at the district level

The omitted cohort is women born between 1954 and 1963. The average age at marriage in around 19 years old, thus women born in the 1964-1973 cohort married around 1983-1992 and women born in the 1974-1983 married approximately during 1993-2002.

8 Appendix

8.1 Proof of Proposition 1

Proposition 1 establishes two testable results regarding the choice between love and arranged marriages. Children calculate their expected utility and compare it for each type of marriage.

$$\begin{aligned} E[u(c_k)] &= \left[1+(1-\varphi)\left(\frac{x_{k,h}\lambda^*_{k,h}+x_{s,h}\lambda_{s,h}}{2}\right)\right]-\frac{d_k}{2}\left[1+(1-\varphi)\left(\frac{x_{k,h}\lambda^*_{k,h}+x_{s,h}\lambda_{s,h}}{2}\right)\right]^2 \\ &-\frac{d_k}{2}\left\{(1-\varphi)^2\sigma^2_\delta\left(\frac{1+\varrho_{ks}\left(I(e=1),\,I(L=1)\right)}{2}\right)\right\}+E\left(\alpha\right) \end{aligned} \quad (27)$$

(i) The assumption on assortative mating allows me to simplify equation 7, in equilibrium $x_{k,h}=x_{s,h}$, which implies that $\lambda_{k,h}=\lambda_{s,h}$. The first term of this equation is positive as long as $x_L-x_A>0$ and $\frac{2p}{\beta}\frac{[1+d(e_{high}-1)]}{(1-d)}<\varphi x_{k,h}<\frac{4p}{\beta}\frac{[1+d(e_{high}-1)]}{(1-d)}$, $h=A,L$:

$$(1-\varphi)\left(x_L\lambda^*_L-x_A\lambda^*_A\right)\left[(1-d_k)-\frac{d_k}{2}(1-\varphi)\left(x_L\lambda^*_L+x_A\lambda^*_A\right)\right]>0 \quad (28)$$

This follows from $\frac{\partial\lambda}{\partial x}=\frac{\beta\varphi(1-d)\left[2dp^2+d\beta\varphi^2x^2\right]-[(1-d)(\beta\varphi x-2p)-2pde]\left[2d\beta\varphi^2x\right]}{d^2(2p^2+\beta\varphi^2x^2)^2}>0$ and $\lambda>0$ as long as the returns to school belong to the range defined above. For simplicity, I am assuming that $e=0$ and $d=d_k$. Therefore, $(x_L\lambda^*_L-x_A\lambda^*_A)>0$ since $x_L>x_A\rightarrow\lambda^*_L>\lambda^*_A$.

As long as the returns to schooling are larger than the lower bound, the increase on x_L-x_A leads to an increase in the probability of love marriage. The remaining question is whether equation 28 holds when x_L reaches the upper limit. Notice that as x_L or x_A increase, the negative term of equation 28 increase as well, $-\frac{d_k}{2}(1-\varphi)\left(x_L\lambda^*_L+x_A\lambda^*_A\right)$. We should consider the possibility that this term becomes larger than $(1-d_k)$, leading children to prefer arranged marriages. I will consider the upper bound of the term $(x_L\lambda^*_L+x_A\lambda^*_A)$, which will be reached when $x_L=x_A=\frac{4p}{\varphi\beta}$.

If both $x_L=x_A=\frac{4p}{\varphi\beta}$, the first part $(x_L\lambda^*_L-x_A\lambda^*_A)$ becomes zero. I am disregarding this effect since I am interested in showing only that:

$$\mid-\frac{d_k}{2}(1-\varphi)\left(x_L\lambda^*_L+x_A\lambda^*_A\right)\mid\leq\mid(1-d_k)\mid \quad (29)$$

when the term on the left hand side reaches the maximum. Therefore, by assuming $x_L=x_A=\frac{4p}{\varphi\beta}$, I can show that **??** holds as long as $1>\frac{(1-\varphi)}{\varphi}\left[\frac{4}{\beta+16}\right]$. Then as long as parents receive a sufficiently large share φ that satisfies this condition, even when the returns to education are close to their upper boundary, the second term will remain be positive. Then, for a given x_A, as $\uparrow x_L$, the gain represented by the term 28 will increase relative to the dis-utility generated by the loss of insurance (arranged marriage).

(ii) Follows from equations 6 and 7. As the potential shock is reduced ($\downarrow\sigma^2_{i\delta}$), the insurance benefits decrease and the utility loss from incurring the effort cost increases, leading parents to switch to low effort (increasing education for the child and, therefore, increasing her outside option). Holding constant education, it also increases $E[u(c_{k,L})+\alpha_L]-E[u(c_{k,A})+\alpha_A]$.

The other parameters of the model also matter for the final decision since they will determine the value of insurance, the investment in education and the decision on effort:

(i) *As*$\varrho_{ks}\left(e,\,I(L=1)\right)-\varrho_{ks}\left(e,\,I(L=0\right)$ *converges to zero. Arranged marriage partners lose their insurance advantage relative to love marriage partners when there is no difference in the dispersion of income between both types of marriages.*

(ii) $d_k>0$ *or* $d>0$ *decreases. More risk-averse agents will prefer arranged marriages over love marriages.*

(iii) $e_{high}>0$ *increases. Parents face a trade-off between exerting high effort and investing in education/consuming; the rising cost of effort will increase the foregone consumption in both periods.*

(iv) $E\left(\alpha_L\right)-E\left(\alpha_A\right)>0$ *when the average in partner "compatibility" is larger in love marriages than in arranged marriages.*

These additional results follow from:

(i) Follows directly from equation 7, as $\varrho_{ks}\left(e,\,I(L=1)\right)-\varrho_{ks}\left(e,\,I(L=0\right)\rightarrow 0$, then $E[u(c_{k,L})+\alpha_L]-E[u(c_{k,A})+\alpha_A]$ increases.

(ii) More risk-averse children will give higher weight to the insurance gain:

$$\frac{\partial(E[u(c_k)]_L-E[u(c_k)]_A)}{\partial d_k}=-\frac{1}{2}\left[1+(1-\varphi)\,x_L\lambda_L^*\right]+\frac{1}{2}\left[1+(1-\varphi)\,x_A\lambda_A^*\right]-\frac{(1-\varphi)^2}{4}\left[\varrho_{ks}\left(e,\,I(L=1)\right)-\varrho_{ks}\left(e,\,I(L=0)\right]\right.$$

where $x_L>x_A\rightarrow\lambda_L^*>\lambda_A^*\rightarrow-\frac{1}{2}\left[1+(1-\varphi)\,x_L\lambda_L^*\right]+\frac{1}{2}\left[1+(1-\varphi)\,x_A\lambda_A^*\right]<0$, and $\varrho_{ks}\left(e,\,I(L=0\right)<0$ by assumption, therefore $-\frac{(1-\varphi)^2}{4}\left[\varrho_{ks}\left(e,\,I(L=1)\right)-\varrho_{ks}\left(e,\,I(L=0)\right]\right.<0$, leading to $\frac{\partial(E[u(c_k)]_L-E[u(c_k)]_A)}{\partial d_k}<0$. More risk-averse children prefer the insurance provided by the arranged marriage.

In the case of parents, a similar result follows:

$$\frac{\partial(E[u(c_k)]_L-E[u(c_k)]_A)}{\partial d}=\left[(1-\varphi)\,x_L\left(1+\lambda_L\right)-d_k\right]\frac{\partial\lambda_L}{\partial d}-\left[(1-\varphi)\,x_A\left(1+\lambda_A\right)-d_k\right]\frac{\partial\lambda_A}{\partial d}$$

where $\frac{\partial\lambda_i}{\partial d}=-\frac{(\beta\varphi x-2p)}{d^2[2p^2+\beta\varphi^2x^2]^2}-\frac{2pe}{d^2[2p^2+\beta\varphi^2x^2]^2}-\frac{[(\beta\varphi x-2p)(1-d)-2pde]\left[2p^2+\beta\varphi^2x^2\right]}{d^2[2p^2+\beta\varphi^2x^2]^2}<0$, in addition $x_L>x_A\rightarrow\lambda_L^*>\lambda_A^*\rightarrow\left[(1-\varphi)\,x_L\left(1+\lambda_L\right)-d_k\right]>\left[(1-\varphi)\,x_A\left(1+\lambda_A\right)-d_k\right]$ and $\frac{\partial^2\lambda}{\partial d\partial x}=\frac{-\beta\varphi\left(2p^2-\beta\varphi^2x^2+4\varphi xp\right)}{d^2(2p^2+\beta\varphi^2)}<0$, therefore $\frac{\partial(E[u(c_k)]_L-E[u(c_k)]_A)}{\partial d}<0$. More risk-averse parents also prefer insurance. They invest in lower education for their child, reducing her outside option and effectively increasing the probability that the child will accept the arranged marriage.

(iii) The derivative of equation 6 with respect to e_H:

$$\left[p\frac{\partial\lambda}{\partial e_H}+1\right]\left[d\left(1-p\lambda^*-e_H\right)-1\right]+\beta\varphi x\frac{\partial\lambda}{\partial e_H}\left[1-d\left(1+\varphi x\lambda^*\right)\right]<0$$

where $\left[p\frac{\partial\lambda}{\partial e_H}+1\right]=1-\frac{2p^2}{2p^2+\beta\varphi^2x^2}>0$; $\left[d\left(1-p\lambda^*-e_H\right)-1\right]<0$ since $\left(1-p\lambda^*-e_H\right)\leq 1$; $\frac{\partial\lambda}{\partial e_H}<0$ and $\left[1-d\left(1+\varphi x\lambda^*\right)\right]=\left(2p^2+2p\varphi x\right)(1-d)+2pde\varphi x>0$. The utility from choosing high effort decreases as the cost of effort increases to the point where parents will switch to low effort, increasing the education of children and their outside option.

Therefore, as e_{high} increases, education decreases $\frac{\partial\lambda(e_{high})}{\partial e_{high}} < 0$ and $\frac{\partial E[u(c_p)]}{\partial e_{high}} < 0$ leading parents to switch to low effort instead.

(iv) Also follows from equation 7; a sufficient condition for *(iv)* is that $F_L(\alpha)$ first order stochastically dominates $F_A(\alpha)$ (by definition of FSD).

8.2 Extending the Model to Include Divorce

Expression 8 might be reduced to[47]:

$$\alpha_h - \beta\left[E\left(\alpha_L\right) - \alpha_h\right] < -\phi - \frac{d_k\beta}{2}\frac{(1-\varphi)^2}{4}\sigma_\delta^2\left[\varrho_{ks}\left(e,\,L=1\right) - \varrho_{ks}\left(e,\,h\right)\right] + \frac{d_k}{2}\left(1-\varphi\right)^2\sigma_\delta^2\left[\frac{\varrho_{ks}\left(e,\,h\right)}{2} - \frac{1}{2}\right] \quad (30)$$

where $\varrho_{ks}\left(e,h\right)$ is the correlation between the child and her spouse and depends on the type of marriage and the effort of parents in the first period; $\varrho_{ks}\left(e,\,L=1\right)$ is the correlation between spouses' income in a love marriage (independent of effort) and σ_δ^2 is the variance of the shock.

The final expression of 30 depends on the type of marriage chosen in period 2 and determines the thresholds for divorce:

(i) If the child chooses love marriage in period 2, then $\varrho_{ks}\left(e,\,L=1\right) = \varrho_{ks}\left(e,h\right)$ and the divorce threshold is given by:

$$\alpha_L < (1+\beta)^{-1}\left[-\phi + \frac{d_k}{2}\left(1-\varphi\right)^2\sigma_\delta^2\left[\frac{\varrho_{ks}\left(e,\,I(L=1)\right)}{2} - \frac{1}{2}\right] + \beta E\left(\alpha_L\right)\right] \quad (31)$$

(ii) If the child chooses arranged marriage in period 2, then $\varrho_{ks}\left(e,\,L=1\right) > \varrho_{ks}\left(e,h\right)$ and the threshold is given by:

$$\alpha_A < (1+\beta)^{-1}\left[-\phi - \frac{d_k\beta}{2}\frac{(1-\varphi)^2}{4}\sigma_\delta^2\left[\varrho_{ks}\left(e,\,I(L=1)\right) - \varrho_{ks}\left(e,\,I(L=0\right)\right] + \frac{d_k}{2}\left(1-\varphi\right)^2\sigma_\delta^2\left[\frac{\varrho_{ks}\left(e,\,I(L=0)\right)}{2} - \frac{1}{2}\right] + \beta E\left(\alpha_A\right)\right] \quad (32)$$

These thresholds are used to derive the divorce probabilities in 9. The solution of the model, therefore, is found by calculating the expected utility from period 2, taking into account these (endogenous) probabilities of divorce for period 3. For a given level of education λ_k, the child will prefer love marriage if:

$$E\left[u\left(c_k\right)\right]^{M,L} + \beta(1-P^{D,L})E\left[u\left(c_k\right)\right]^{M,L} + \beta P^{D,L}E\left[u\left(c_k\right)\right]^{D,L} > E\left[u\left(c_k\right)\right]^{M,A} + \beta(1-P^{D,A})E\left[u\left(c_k\right)\right]^{M,A} + \beta P^{D,A}E\left[u\left(c_k\right)\right]^{D,A} \quad (33)$$

[47] I am assuming that individuals in arranged marriages will find a partner with returns $x_A$in the love marriage market. If this assumption is relaxed, the threshold for divorce for individuals in arranged marriages will depend directly on the difference of returns $x_L - x_A$and on the education level λ:

$$\begin{aligned}\alpha_A \quad < \quad & (1+\beta)^{-1}\left\{-\phi - \frac{d_k\beta}{2}\frac{(1-\varphi)^2}{4}\sigma_\delta^2\left[\varrho_{ks}\left(e,\,L=1\right) - \varrho_{ks}\left(e,\,L{=}0\right)\right] + \beta(1-\varphi)\lambda\left(x_L - x_A\right)\right.\\ & \left.\left[1 - d_k - \frac{d_k}{2}(1-\varphi)\left(x_L + x_A\right)\lambda\right] + \frac{d_k}{2}\left(1-\varphi\right)^2\sigma_\delta^2\left[\frac{\varrho_{ks}\left(e,\,L{=}0\right)}{2} - \frac{1}{2}\right] + \beta E\left(\alpha_L\right)\right\}\end{aligned}$$

If this is the case, the optimal solution for λ^*in period 1 depends on the partial derivative of the probability of divorce with respect to λ ($\frac{\partial P^d}{\partial\lambda}$) delivering a cubic term on λ. There is no closed form solution for this case and we need to rely on numerical solutions. For simplicity, I am assuming that this is not the case. However, if the assumption is relaxed, it will deliver a higher divorce rate for arranged marriages as $x_L$increases relative to x_A. This in turn is internalized by parents in period 1 when choosing education and effort.

In period 1, parents will choose the education and effort levels also taking into account the probability of divorce:

$$\lambda(e) = \frac{(1-d)\left\{(1-\varphi)\left[\left(\beta+\beta^2\left(1-P^D\right)\right)(1+\beta)+2\beta^2P^D\right]x+(1-\varphi)\,\beta^3P^Dx-2p\right\}-2dep}{d\left\{2p^2+(1-\varphi)^2\left[\left(\beta+\beta^2\left(1-P^D\right)\right)(1+\beta)+2\beta^2P^D\right]x+(1-\varphi)^2\beta^3P^Dx\right\}} \tag{34}$$

These results show that individuals will divorce depending on the realization of the love term relative to the economic characteristics of the spouse (her insurance quality)

Proposition 4. *For both type of marriages, the probability of divorce will:*

(i) *Increase if the discounted expected match quality from a new draw from the love distribution is larger than that of the current partner:* $\uparrow\left[\beta E\left(\alpha_L\right)(1+\beta)^{-1}-\alpha_h,\right]\ h=L,A.$

(ii) *Increase if the income covariance between spouses increases for a given size (variance,*σ_δ^2*) of the shock.*

(iii) *Decrease if the size of the shock (*σ_δ^2*) increases for a given income covariance between spouses.*

(iv) *Decrease as the cost of divorce increases,* $\uparrow\phi$*.*

Proof of proposition 4. They follow directly from expressions 31 and 32 combined with equation 9.

8.3 Proposition 3 and Numerical Comparative Statics with Different Gender Composition:

Proposition 3. *If* $\varrho_{g_j,b_n}(e_{g,j}=1,e_{b,n}=1)\Rightarrow\varrho_{k,s}=1\ k=g_j,b_n,\ s=s_j,s_n$ *(the most constrained case, only one potential insurance partner, each component of* $\varrho_{g_1,b_1}(e_{g,j},e_{b,i})$ *has a positive correlation equal to 1) and* $\varrho_{g_j,b_i}(e_{g,j}=0,e_{b,i}=1)=\varrho_{g_j,b_i}(e_{g,j}=0,e_{b,i}=0)=0$*, then ceteris paribus:*

a) If $g=1/2$*, families are composed of one boy and one girl, and if* $x_b^h/p_b^h > x_g^h/p_g^h$*,* $h=A,L$ *(Strauss and Thomas, 1996; Behrman, 1997), the optimal education level is: (i)* $\lambda_b>0$ *for the boy; and (ii)* $\lambda_g=0$ *for the girl. Given the choice of education, parents endogenously decide to exert high effort for the girl* $e_{g,1}=1$ *and low effort for the boy* $e_{b,1}=0$ *(given a low enough love term for the girl, such that she does not reject the arranged marriage). The education of the boy endogenously responds to his marriage decision in the second period: (i)* $\lambda_b(x_L)$ *if he chooses love marriage with returns* x_L*; or (ii)* $\lambda_b(x_A)$ *if he chooses the proposed arranged marriage (corresponding to the low insurance quality mate) with returns* x_A*. And if* $x_L>x_A$*, then* $\lambda_b(x_L)>\lambda_b(x_A)$*.*

b) If $g=1$ *or* $g=0$*, families are composed of two girls or two boys, and if they are identical in* $p_g(p_b)$ *and* $x_g(x_b)$*, then parents toss a coin and offer with 50% probability the high insurance quality mate to girl (boy) 1 (*$e_{g,1}=1$*) and the low insurance quality mate to girl (boy) 2 (*$e_{g,2}=0$*), conditional on the high insurance quality arranged marriage being accepted. The education level of both girls (boys) responds endogenously to the marriage decision of the second girl (boy). If she (he) decides to reject the low quality partner and* $x_L>x_A$*, then* $\lambda_{2g}(x_{gL})>0$*,* $\lambda_{1g}(x_{gA})=0$*. If the girl (boy) decides to accept the low quality arranged marriage, then* $\lambda_{2g}(x_{gA})=\lambda_{1g}(x_{gA})>0$*.*

Figure 13, panel A shows comparative statics for the game with one son and one daughter. It is important to emphasize that although the decision of each child depends on equation 7, the second period utility for each of them depends on the set of strategies of the three agents (parents, son and daughter). The agents affect each other through the budget constraint (education and effort are costly) and through $x_b^h/p_b^h \gtrless x_g^h/p_g^h$. The first graph in panel A, column 1 numerically simulates the model and shows that as the net returns for boys increase relative to the net returns for girls ($\left(\frac{x_b^L}{p_b} - \frac{x_b^A}{p_b}\right) - \left(\frac{x_g^L}{p_g} - \frac{x_g^A}{p_g}\right)$), the probability that the son has an arranged marriage decreases and the probability that the daughter accepts the arranged marriage increases. Column 2 shows that parents exert high effort for the child with the lowest net returns to schooling (the effort decreases for boys and increases for girls as the returns shift in favor of boys). And finally, column 3 shows that the results are achieved through differential investments in education for boys and girls.

Panel B of figure 13 shows the numerical comparisons for the game with two girls. The first graph plots the probability of arranged marriage (high effort) of girl 1 against the probability of arranged marriage (high effort) of girl 2. Since by construction parents have incentives to marry only one of them ($\varrho_{g_j,b_n}(e_{g,j} = 1, e_{b,n} = 1) \Rightarrow \varrho_{k,s} = 1\, k = g_j, b_n,\ s = s_j, s_n$ and $\varrho_{g_1,g_2}(e_{g_1} = 0, e_{g_2} = 1) = 0$), their probabilities (of having an arranged marriage) are inversely correlated. The second graph shows the response of the arranged marriage probability as $(x_{gL} - x_{gA}) \uparrow$; since both girls are identical, both of them face the same trade-off between insurance and returns outside the network, leading to a decreasing probability of arranged marriage for each of them. The last graph plots the probability of both girls receiving positive education. When both girls choose the same type of marriage, the model delivers identical education for them, illustrated by a positive relation between each pair of variables in this graph.

8.3.1 Proof of Proposition 3:

Proposition 3 compares two-children households with different gender composition under the assumption that boys and girls might have different prices/returns to education:

a) The proof of the education levels follows directly from the results of equation 13. Since the returns (per dollar spent) for the boy are larger than the returns (per dollar spent) for the girl, parents choose to educate only the boy. Parents still have incentives to acquire insurance since:

$$E\left[u(c_f\left(\lambda_{b,1} > 0, \lambda_{g,1} = 0, e_{g,1} = 1, e_{b,1} = 0\right)\right] = \left[1 + \varphi\left(\frac{1}{2}x_b\lambda_b^*\right)\right] - \frac{d}{2}\left[1 + \varphi\left(\frac{1}{2}x_b\lambda_b^*\right)\right]^2 - \frac{d}{4}\sigma_\delta^2\varphi^2 >$$

$$E\left[u(c_f\left(\lambda_{b,1} > 0, \lambda_{g,1} = 0, e_{g,1} = 0, e_{b,1} = 0\right)\right] = \left[1 + \varphi\left(\frac{1}{2}x_b\lambda_b^*\right)\right] - \frac{d}{2}\left[1 + \varphi\left(\frac{1}{2}x_b\lambda_b^*\right)\right]^2 - \frac{d}{2}\sigma_\delta^2\varphi^2$$

Therefore parents will prefer to offer the arranged marriage to the girl since her outside option is low enough ($\lambda_g = 0$) such that she will accept the arranged marriage as long as $E(\alpha_L) - E(\alpha_A)$ is sufficiently small (recall that each child still decides the type of marriage based on equation 7). The numerical comparative statics of panel A of figure 13 provide a graphical analysis for this case.

(b) If the two children are identical, parents are indifferent choosing between them for the arranged marriage. Parents will calculate the expected utility under each scenario and choose education and effort that gives them the highest expected utility (payoff):

$$E[u(c_f(\lambda_{g1}(e_{g1}),\lambda_{g2}(e_{g2}),e_{g1} = 1,e_{g2}=0)] = E[u(c_f(\lambda_{g1}(e_{g1}),\lambda_{g2}(e_{g2}),e_{g1}=0,e_{g2}=1)] =$$
$$\left[1+\frac{\varphi}{2}\left(x_{g1}\lambda_{g1}(e_{g1})+x_{g2}\lambda_{g2}(e_{g2})\right)\right] - \frac{d}{2}\left[1+\frac{\varphi}{2}\left(x_{g1}\lambda_{g1}(e_{g1})+x_{g2}\lambda_{g2}(e_{g2})\right)\right]^2 - \frac{d}{4}\sigma_\delta^2\varphi^2 \tag{35}$$

$$E[u(c_f(\lambda_{g1}(e_{g1}),\lambda_{g2}(e_{g2}),e_{g1} = 0,e_{g2}=0)] =$$
$$\left[1+\frac{\varphi}{2}\left(x_{g1}\lambda_{g1}(e_{g1})+x_{g2}\lambda_{g2}(e_{g2})\right)\right] - \frac{d}{2}\left[1+\frac{\varphi}{2}\left(x_{g1}\lambda_{g1}(e_{g1})+x_{g2}\lambda_{g2}(e_{g2})\right)\right]^2 - \frac{d}{2}\sigma_\delta^2\varphi^2 \tag{36}$$

Then equation 35 is larger than equation 36 for sufficiently low (unconstrained) returns (since both children are identical, and the unconstrained returns are high enough, they will prefer to educate both children and give up the insurance). Parents will choose to offer the arranged marriage to girl 1 (girl 2) if $E(\alpha_{g1,L}) - E(\alpha_{g1,A})$ is sufficiently small (equivalent to girl 1 (girl 2), accepting the arranged marriage with the high insurance quality groom).

The education levels depend on the choices of the girl offered the low insurance quality arranged marriage ($e_{g,j} = 0$). From the maximization problem 11, if she accepts the low quality arranged marriage, then $x_{g1,A} = x_{g2,A}$, and since both girls face p_g cost of education, then $\lambda_{2g}(x_{gA}) = \lambda_{1g}(x_{gA}) > 0$. If she rejects the low insurance quality arranged marriage, and if $x_{g,L} > x_{g,A}$, then from equations 12 and 13 we have a corner solution where the girl with the higher returns receives all the education, $\lambda_{g,2}(x_{g,L}) > 0$, $\lambda_{g,1}(x_{g,A}) = 0$. The numerical comparative statics of panel B of figure 13 provide a graphical analysis to the solution of this problem when the two children are identical.

8.4 Size of the network and the number of children

This section shows how the number of children affects the role of the size and insurance quality of the network. For the rest of this section, I assume that insurance quality of the network refers to income covariance across the households where the children are married (conditional on having an arranged marriage). For simplicity, I abstract from the children's decision. I only consider the side of the parents who face a shock; they have incentives to smooth it out by marrying off their N_k children with their insurance partners. If we consider the children as decision makers, the analysis becomes increasingly complicated as the number of players in the game increases to $N_k + 1$.

Let ε be the shock faced by parents in period 2 with mean 0 and variance σ_ε^2. If the parents marry off all their children, they pool their resources with the N_k households and consume the average. Their consumption in the second period is given by:

$$c_f = 1 + \frac{\varepsilon + \sum_i^{N_k}\varepsilon_i}{N_k+1}$$

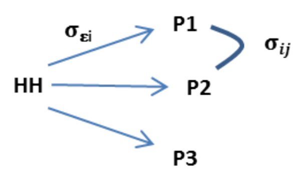
The size of the network (including the household) is $N_k + 1$, corresponding to the total number of children plus the parents. The parents calculate their expected utility in period 2:

$$E[u(c_f)]_A = E\left\{\left(1+\frac{\varepsilon+\sum_i^{N_k}\varepsilon_i}{N_k+1}\right)-\frac{d}{2}\left(1+\frac{\varepsilon+\sum_i^{N_k}\varepsilon_i}{N_k+1}\right)^2\right\}$$

$$E[u(c_f)]_A = 1-\frac{d}{2}\left\{1+\frac{\sigma_\varepsilon^2}{(N_k+1)^2}+\frac{1}{(N_k+1)^2}\left[\sum^{N_k}\sigma_i^2+\sum^{m}\sigma_{ij}+\sum^{N_k}\sigma_{\varepsilon i}\right]\right\},\ m=\begin{pmatrix} N_k \\ 2 \end{pmatrix}=\frac{N_k!}{2!(N_k-2)!} \qquad (37)$$

where σ_ε^2 is the variance of the shock of the parents' household, σ_i^2 is the variance of the N_k households where children are married, σ_{ij} is the covariance across the households where children are married, and $\sigma_{\varepsilon i}$ is the covariance between the parents' household and the other households.

From expression 37, the need of a large and high quality insurance network becomes evident. As the number of children increases, so does the number of income covariances between them:



The total number of covariances that parents should be consider is: $m=\begin{pmatrix} N_k \\ 2 \end{pmatrix}=\frac{N_k!}{2!(N_k-2)!}$. Households belonging to a small social network face a potentially large dis-utility if they arranged the marriage of all their children.

8.5 Data Appendix

This paper uses data from several sources in order to show changes in insurance motive. The challenge is to show that wages, income, income variance and education varied in response to the program at the time (or before) individuals made their marital choices. In order to show it, I have used data from the Population Census and the Socioeconomic Surveys (SUSENAS) which were conducted approximately at the same time that the Bimas/Inmas program was delivering successful outcomes. The inter-census from 1976 is the first survey that contains any information on income.

The data used to construct the intensity of the Green Revolution and measure its impact was collected from several printed sources. The main variables used for the analysis come from the 1963 and 1983 agricultural census. The data on production and land utilization come from Kompilasi data reports (reports on production by district in 1980, 1981 and 1982) and from Luas tanah menurut penggunaannya Jawa-Madura & di luar Jawa (land utilization reports from Java and Outer Java).

The analysis in table 3 was done using information on wages from the 1976 and 1995 inter-census surveys, and the 1980, 1981, and 1982 SUSENAS. I restricted the sample to individuals aged 17 to 65 in each survey year, with non-missing information on education, wages, district of residence and the other relevant demographic characteristics. Panels A and B of table 16 present the average characteristics and standard deviation for each of the relevant variables used in the analysis, as well as the total number of observations and source of the data. Panel C reports the same information for the 1993 IFLS sample of wage earners matched with districts with information on the intensity of the Bimas/Inmas program in 1982/83 and for the sample matched with the information on agricultural characteristics from the 1963 census (Note that some districts in 1983 did not appear in the 1963 agricultural census, and that accounts for the difference in observations reported in table 16; the results, however, do not change substantially).

The information used for panel A of table 4 comes from the 1981, 1982, 1984 and 1987 SUSENAS surveys. I used information on labor income from all sources at the household level: wages and profits (agricultural and non-agricultural). The measure I used refers to the total household income in real terms divided by the total number of household members, and I focused on households that report agriculture as the main source of income. Table 17 summarizes the per capita figures by year in dollars using the 1993 PPP exchange rate. It also show the total number of observations by year and reports the information for agricultural and non-agricultural households. Panel B of table 17 summarizes the information of the 1993 IFLS sample used for the analysis of panel B in table 4. As with the analysis in panel A, I use only information on labor income (from all sources). I present the summary statistics for the districts matched with the 1983 Bimas/Inmas intensity, and separately for the districts matched with the 1963 agricultural characteristics.

Finally, the analysis in table 6 used data from the 1995 inter-census survey focusing on individuals born before 1978 and matched to their district of birth. The final sample used was determined based on having information on the Bimas/Inmas program, the 1963 agricultural characteristics and the data provided by Duflo (2001) on the intensity of the school construction program (1973-1978).

8.6 Additional Graphs and Tables

Table 9: Test of Full Insurance: Robustness Checks

	First Differences					
	Alternative Expenditure		Robust Regression		Dropping 1% top/botom	
	Arranged	Love	Arranged	Love	Arranged	Love
Log Agg Cons	0.498***	0.314***	0.471***	0.347***	0.482***	0.334***
	(0.0870)	(0.0525)	(0.0753)	(0.0382)	(0.0909)	(0.0529)
Log Income	**-0.000995**	**0.00690***	**-0.00172**	**0.00918*****	**-0.00135**	**0.0259*****
	(0.00572)	(0.00403)	(0.00499)	(0.00256)	(0.00945)	(0.00605)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	820	2,876	820	2,876	790	2,776
R-squared	0.054	0.074	0.085	0.093	0.050	0.091
Joint Test (Log Agg Cons = 1) (Log Income = 0)						
F-statistic	16.67	86.57	24.95	154.30	16.28	87.71
P-value	0.00	0.00	0.00	0.00	0.00	0.00

	First Differences					
	Wages+Profits		1993-1997		1997-2000	
	Arranged	Love	Arranged	Love	Arranged	Love
Log Agg Cons	0.470***	0.341***	0.403***	0.261***	0.472***	0.367***
	(0.0855)	(0.0512)	(0.0986)	(0.0607)	(0.129)	(0.0671)
Log Income	**-0.00419**	**0.00846****	**-0.00126**	**0.0107****	**-0.00717**	**0.00754***
	(0.00484)	(0.00367)	(0.00479)	(0.00517)	(0.00950)	(0.00438)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	820	2,876	410	1,438	410	1,438
R-squared	0.051	0.087	0.070	0.085	0.036	0.090
Joint Test (Log Agg Cons = 1) (Log Income = 0)						
F-statistic	19.67	85.14	18.32	76.10	8.89	47.01
P-value	0.00	0.00	0.00	0.00	0.00	0.00

All results are for a first differences model. Column 1 adds education and medical expenditure to the consumption expenditure; column 2 uses a robust regression which gives lower weight to observations that are potential outliers; column 3 drops the bottom and top 1% of the income distribution in levels by year; column 4 uses only wages and profits as income; column 5 uses period 1993 and 1997; and finally column 6 uses period 1997 and 2000. Standard errors clustered at the household level.

Table 10: Test of Full Insurance: IV using lagged changes in log aggregate consumption and lagged changes in log income

First Differences: IV using lagged consumption and lagged income	Arranged		Love	
	OLS	IV	OLS	IV
Log Agg Cons	0.472***	0.505**	0.367***	-0.107
	(0.129)	(0.212)	(0.0671)	(0.122)
Log Income	**-0.00717**	**-0.00635**	**0.00754***	**0.00821**
	(0.00950)	(0.0172)	(0.00438)	(0.0122)
Controls	Yes	Yes	Yes	Yes
F-statistic		104.23		117.35
		9.88		16.2
Observations	410	410	1,438	1,438
R-squared	0.036	0.036	0.090	0.050

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

First Stage: IV using lagged consumption and lagged income	Arranged		Love	
	Log Agg Cons 2000-1997	Log Inc 2000-1997	Log Agg Cons 2000-1997	Log Inc 2000-1997
Log Agg Cons 1997-1993	-0.458***	-0.716	-0.479***	0.495
	(0.0325)	(0.585)	(0.0314)	(0.364)
Log Income 1997-1993	-0.00276	-0.479***	0.00136	-0.429***
	(0.00209)	(0.116)	(0.00139)	(0.0765)
Controls	Yes	Yes	Yes	Yes
Observations	410	410	1,438	1,438
R-squared	0.354	0.276	0.277	0.152

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The change in log aggregate consumption between 1997 and 2000 is instrumented using the change in log aggregate consumption between 1993 and 1997. Similarly, the change in log household income between 1997 and 2000 is instrumented using the change in log household income between 1993 and 1997. Standard errors clustered at the household level.

Table 11: 1963 Agricultural Characteristics and 1982/83 Bimas/Inmas Program Characteristics

1963 Agricultural Characteristics by District				
Agricultural Charact.	Mean	Std. Dev.	Min	Max
% Sawah Land	36%	21%	0%	92%
# Farms	62,396	53,720	2,737	513,274
Total Agric Land (Ha)	66,259	62,278	1,858	499,280
Farms/Ha	1.09	0.53	0.13	2.85

BIMAS Program in 1982/1983 by District					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
% Area Bimas/Inmas	204	70%	32%	0%	100%
% Total Irrigated Land	205	24%	20%	0%	85%
% with Irrigation System	205	16%	17%	0%	83%
% Rainfed and other irrigation	205	8%	8%	0%	51%
% HH using Fert 1982/1983	204	77%	29%	1%	100%
Avg use of urea (kg)	198	134.4	93.8	0.0	333.4
Avg use of tsp (kg)	198	52.1	39.9	0.0	157.5
Avg use of kcl (kg)	198	5.0	5.7	0.0	31.9
% Land with VUTW I Seeds	204	12%	12%	0%	100%
% Land with VUTW II Seeds	204	53%	30%	0%	95%
% Land with VUB Seeds	204	9%	10%	0%	72%
% Land with VUN Seeds	204	1%	2%	0%	20%
% Land with Local Seeds	204	26%	31%	0%	100%

Table 12: Correlation of Intensity of Bimas/Inmas Program and Agricultural Inputs in 1982/83

Area Bimas/Inmas and Agricultural Inputs in 1982/1983				
	% Land with Irrigation System	Avg use of urea (kg)	% HH using Fertilizer 1982/1983	% Land with VUTW II Seeds
	(1)	(2)	(3)	(4)
% Area Bimas/Inmas	0.334***	257.2***	0.885***	0.718***
	(0.0289)	(10.23)	(0.0163)	(0.0429)
Constant	-0.0721***	-47.80***	0.152***	0.0195
	(0.0223)	(7.983)	(0.0126)	(0.0331)
Observations	203	196	203	203
R-squared	0.400	0.765	0.936	0.582

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Effect of Agricultural Inputs on Wet Paddy Yields

Dep. Var.: Yield of Sawah Paddy					
	(1)	(2)	(3)	(4)	(5)
% Area Bimas/Inmas	1.924***	1.333***	-0.0490	-0.117	
	(0.119)	(0.139)	(0.192)	(0.226)	
% Land with Irrigation System		1.766***	0.656**	0.653**	0.656**
		(0.263)	(0.255)	(0.259)	(0.258)
Avg use of urea (kg)			0.00679***	0.00679***	0.00652***
			(0.000747)	(0.000786)	(0.000592)
% Land with VUTW II Seeds				0.00278	0.00947
				(0.213)	(0.213)
% Land with Local Seeds				-0.0882	-0.0430
				(0.198)	(0.178)
Constant	1.990***	2.117***	2.371***	2.442***	2.379***
	(0.0918)	(0.0853)	(0.0788)	(0.186)	(0.141)
Observations	202	202	195	195	195
R-squared	0.567	0.647	0.751	0.751	0.751

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14: Reduced Form effect of 1963 Agricultural Characteristics on Inputs

Agricultural Inputs in 1980s					
	% Land with Irrigation System	Avg use of urea (kg)	% HH using Fertilizer 1982/1983	% Land with VUTW II Seeds	% Land with Local Seeds
	(1)	(2)	(3)	(4)	(5)
% Sawah Land	-0.00726	101.5	0.666***	0.885***	-0.842***
	(0.116)	(68.90)	(0.216)	(0.241)	(0.261)
Farms/Ha	-0.0377	84.72***	0.407***	0.303***	-0.379***
	(0.0410)	(24.43)	(0.0766)	(0.0852)	(0.0922)
% Sawah Land*Farms/Ha	0.378***	11.52	-0.370*	-0.457**	0.604***
	(0.102)	(60.61)	(0.191)	(0.213)	(0.230)
Constant	0.0348	-1.000	0.256***	0.0965	0.690***
	(0.0378)	(22.93)	(0.0706)	(0.0785)	(0.0850)
Observations	183	177	183	183	183
R-squared	0.486	0.425	0.405	0.257	0.189

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 15: Reduced Form effect of 1963 Agricultural Characteristics on Intensity of Bimas/Inmas and Output

Intensity of Bimas/Inmas and Output 1980s		
	% Area Bimas/Inmas	Yield of Sawah Paddy
	(1)	(2)
% Sawah Land	0.696***	1.355**
	(0.237)	(0.588)
Farms/Ha	0.423***	0.972***
	(0.0839)	(0.208)
% Sawah Land*Farms/Ha	-0.347*	-0.415
	(0.209)	(0.520)
Constant	0.145*	1.967***
	(0.0773)	(0.192)
Observations	183	183
R-squared	0.409	0.451

Standard errors in parentheses

Table 16: Sample Statistics of Wages

Wages Statistics and Demographic Characteristics by Sample								
Panel A. Wages in PPP Dollars by Year								
Year	**Median**	**Mean**	**Std. Dev.**	**Max**	**Min**	**Obs**	**Source**	
1976	50.09	84.22	143.04	7,141	7.16	12,824	Census	
1980	62.56	302.40	2178.10	42,618	0.22	25,089	Susenas	
1981	76.08	124.53	168.26	13,368	0.95	19,805	Susenas	
1982	87.60	120.58	122.81	4,043	3.37	30,571	Susenas	
1995	83.06	104.63	107.77	6,592	6.59	43,504	Census	
Total	77.5	147.0	960.5	42,618	0.2	131,793		
Panel B. Average Characteristics of the Sample by Year								
	Mean				**Standard Deviation**			
Year	**Age**	**Female**	**Urban**	**Education**	**Age**	**Female**	**Urban**	**Education**
1976	34.41	30.8%	27.2%	2.86	11.17	46.2%	44.5%	2.09
1980	34.93	29.1%	27.6%	2.70	11.14	45.4%	44.7%	1.78
1981	34.65	28.8%	28.0%	3.01	11.30	45.3%	44.9%	1.96
1982	35.75	34.4%	29.6%	2.82	11.55	47.5%	45.6%	1.84
1995	34.60	28.5%	28.6%	4.04	10.93	45.1%	45.2%	2.20
Total	34.92	30.2%	28.4%	3.23	11.21	45.9%	45.1%	2.08
Panel C. Wages in PPP Dollars and Average Characteristics of the Sample 1993 IFLS								
	Median	**Mean**	**Std. Dev.**	**Max**	**Min**	**Obs**		
BIMAS/INMAS	67.0	262.7	2769.4	79317.8	0.4	2,730		
Area Sawah	58.2	228.5	2589.8	79317.8	0.4	2,394		
	Mean				**Standard Deviation**			
	Age	**Female**	**Urban**	**Education**	**Age**	**Female**	**Urban**	**Education**
BIMAS/INMAS	39.5	32.1%	37.7%	5.70	12.0	47%	48%	4.68
Area Sawah	39.3	31.0%	44.2%	6.04	11.9	46%	50%	4.76

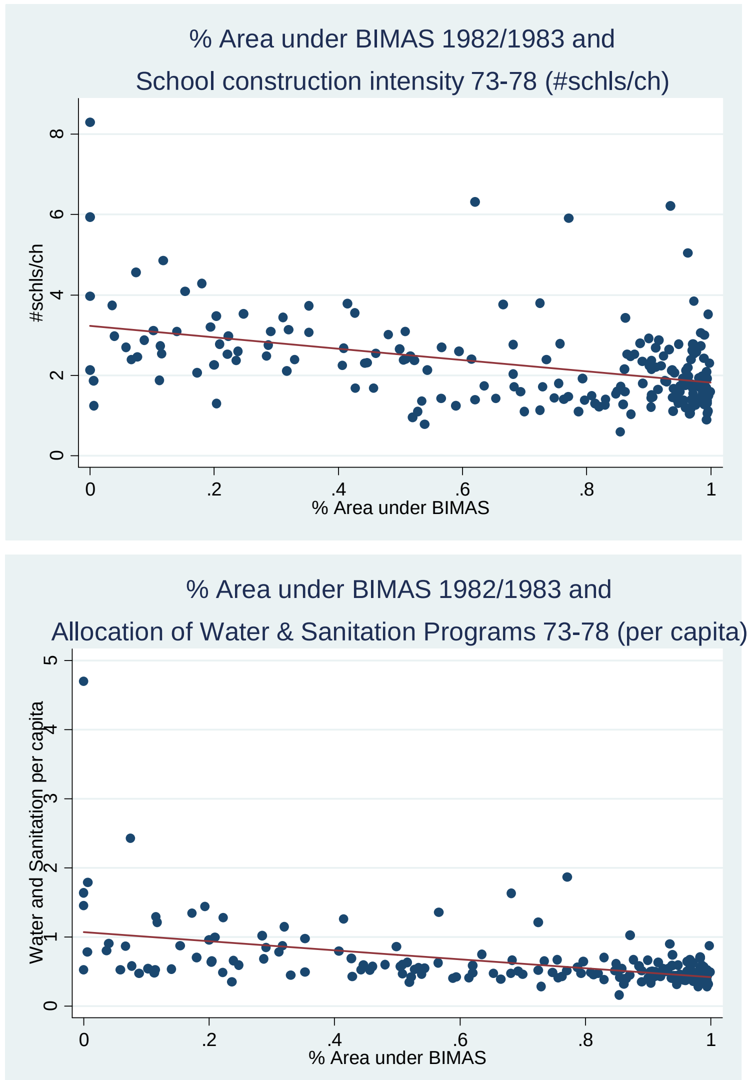
BIMAS/INMAS refers to the sample of IFLS merged with the districts with information on intensity of Bimas/Inmas in 1982/83. Area Sawah refers to the sample of IFLS merged with the districts with information on 1963 agricultural characteristics.

Table 17: Sample Statistics of per capita Income

Per Capita Income Statistics by Sample in PPP Dollars							
Panel A. Per Capita Income in PPP Dollars by Year (SUSENAS)							
Year	**Median**	**Mean**	**Std. Dev.**	**Max**	**Min**	**Obs**	**% Agric HH**
1981	18.64	12.66	85.69	9783.70	0.00	15,112	67.7%
1982	14.36	9.99	57.62	5415.67	0.00	10,399	64.5%
1984	15.52	11.47	23.70	1529.97	0.69	32,778	59.9%
1987	17.01	12.32	31.75	3157.97	0.36	19,049	60.5%
Total	16.34	11.74	48.68	9783.70	0.00	77,338	62.2%
Agricultural Households							
Year	**Median**	**Mean**	**Std. Dev.**	**Max**	**Min**	**Obs**	
1981	16.13	11.57	30.43	1143.14	0.00	10,227	
1982	11.04	7.93	24.47	1359.20	0.00	6,704	
1984	12.33	9.97	14.28	1134.14	0.69	19,619	
1987	12.70	10.49	10.91	435.15	0.97	11,529	
Total	13.05	10.19	19.88	1359.20	0.00	48,079	
Non-agricultural households							
Year	**Median**	**Mean**	**Std. Dev.**	**Max**	**Min**	**Obs**	
1981	23.88	15.31	144.01	9783.70	0.20	4,885	
1982	20.40	14.62	90.57	5415.67	0.00	3,695	
1984	20.27	14.43	32.52	1529.97	0.73	13,159	
1987	23.63	16.76	47.95	3157.97	0.36	7,520	
Total	21.75	15.22	74.61	9783.70	0.00	29,259	
Panel B. Per Capita Income in PPP Dollars 1993 (IFLS)							
	Median	**Mean**	**Std. Dev.**	**Max**	**Min**	**Obs**	
BIMAS/INMAS	30.61	41.73	51.82	1510.84	0.00	4,817	
Area Sawah	32.31	46.70	57.57	1510.84	0.00	5,371	

BIMAS/INMAS refers to the sample of IFLS merged with the districts with information on intensity of Bimas/Inmas in 1982/83. Area Sawah refers to the sample of IFLS merged with the districts with information on 1963 agricultural characteristics.

Figure 12: Intensity of Bimas/Inmas Program and School Construction Program and Allocation of Water and Sanitation



% Area under Bimas/Inmas refers to percentage of agricultural wet paddy (rice) land under any type of intensification in 1982/83. School construction intensity 73-78 and allocation of water & sanitation programs 73-78 were collected and used by Duflo (2001). The former refers to the total number of INPRES schools planned to be constructed between 1973 and 1978 for 1000 children.

Table 18: Effect of the Green Revolution on Returns to Education (controlling for other programs)

Panel A. Effect on Wages using SUSENAS and Census

Dep. Var: ln wage	BIMAS/INMAS		Reduced Form		IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling*BIMAS/INMAS	**0.0293*****	**0.0228*****			**0.0664*****	**0.0484*****
	(0.00698)	(0.00695)			(0.0157)	(0.0150)
Years of schooling*Sawah Area			**0.0614****	**0.0487***		
			(0.0240)	(0.0266)		
Years of schooling*Farms per Ha			**0.0376*****	**0.0316*****		
			(0.00791)	(0.00831)		
Years of schooling*% Sawah Area* Farms/Ha			**-0.0588*****	**-0.0519****		
			(0.0205)	(0.0230)		
Years of schooling	0.0671***	0.0723***	0.0487***	0.0588***	0.0309**	0.0454***
	(0.00734)	(0.00737)	(0.00942)	(0.00914)	(0.0140)	(0.0136)
Female	-0.575***	0.0228***				-0.00240
	(0.0544)	(0.160)	(0.0599)	(0.288)	(0.0712)	(1.200)
District FE	yes	yes	yes	yes	yes	yes
Interaction with Islands Dummy	no	yes	no	yes	no	yes
Observations	131,793	131,793	131,793	131,793	131,793	131,793
R-squared	0.725	0.726	0.725	0.726	0.724	0.726
F-statistic of excluded instruments					16.23	11.42

Panel B. Effect on Wages using IFLS

	BIMAS/INMAS		Reduced Form		IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling*BIMAS/INMAS	**0.0631****	**0.0576****			0.183	0.167
	(0.0282)	(0.0286)			(0.159)	(0.147)
Years of schooling*Sawah Area			**0.0582**	**0.0855**		
			(0.0787)	(0.0728)		
Years of schooling*Farms per Ha			**0.0340**	**0.0716****		
			(0.0332)	(0.0300)		
Years of schooling*% Sawah Area* Farms/Ha			**-0.0543**	**-0.0881**		
			(0.0620)	(0.0585)		
Years of schooling	0.0945***	0.0739*	0.105**	0.0208	-0.0160	-0.0389
	(0.0351)	(0.0379)	(0.0512)	(0.0428)	(0.147)	(0.144)
Constant	2.602***	3.009***	2.501***	0.774*	2.244*	30.99***
	(0.428)	(0.278)	(0.658)	(0.414)	(1.224)	(8.391)
Province FE	yes	no	yes	no	yes	no
District FE	no	yes	no	yes	no	yes
Observations	2,394	2,394	2,442	2,442	2,298	2,298
R-squared	0.376	0.430	0.370	0.439	0.360	0.420
F-statistic of excluded instruments					2.34	2.31

Standard errors clustered at the district level

Panel A. Years of education were imputed for 1980, 1981 and 1982 using self-reported information on highest level of education completed. 1976 and 1995 report the number of years of education. Panel B. The variable on years of education were constructed using information on highest level of education and grade completed (see data appendix for other details in the samples). Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83; Sawah Area = % of agricultural land available for production of wet paddy (rice) in 1963; Farms/Ha = Total number of farms in 1963/ Total Agricultural Land (Ha) in 1963. Additional controls are number of INPRES schools 73-78 per 1000 children, population in 1971 and water & sanitation programs 73-78 per capita all by district and interacted with years of schooling.

Table 19: Effect of the Green Revolution on Mean and Variance of Income (controlling for other programs)

Panel A. Effect on per capita Income using SUSENAS and Census

MLE effect on log pc income mean and variance: Agricultural Households								
	Effect of BIMAS/INMAS				**Reduced Form of Sawah Land and Farms/ha**			
	Income	Variance	Income	Variance	Income	Variance	Income	Variance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Area BIMAS/INMAS	**0.0934**	**0.0466**	**0.107**	**-0.105**				
	(0.0851)	(0.0857)	(0.103)	(0.180)				
% Sawah Area					**-0.0670**	**-0.493**	**0.308**	**-0.988**
					(0.242)	(0.446)	(0.325)	(0.644)
Number farms per Ha					**-0.324*** **	**-0.252**	**-0.217**	**-0.403**
					(0.104)	(0.174)	(0.137)	(0.252)
% Sawah Area* Farms/Ha					**0.547*** **	**0.413**	**0.300**	**0.737**
					(0.201)	(0.345)	(0.268)	(0.483)
Rural	-0.202***	-0.0188	-0.202***	-0.0188	-0.197***	-0.0160	-0.201***	-0.0140
	(0.0225)	(0.0192)	(0.0226)	(0.0203)	(0.0219)	(0.0196)	(0.0218)	(0.0203)
Constant	10.27***	0.765***	10.25***	0.891***	10.44***	1.045***	10.30***	1.262***
	(0.0821)	(0.0863)	(0.0884)	(0.175)	(0.130)	(0.230)	(0.158)	(0.346)
Province FE	yes	yes	yes	yes	yes	yes	yes	yes
Interaction with Islands Dummy	no	no	yes	yes	no	no	yes	yes
Observations	48,079	48,079	48,079	48,079	48,079	48,079	48,079	48,079
Wald chi2								

Panel B. Effect on per capita Income using IFLS

	Income	**Variance**	**Income**	**Variance**	**Income**	**Variance**	**Income**	**Variance**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Area BIMAS/INMAS	**0.988** **	**-1.380**	**1.330** **	**-0.777**				
	(0.487)	(1.825)	(0.573)	(1.357)				
Age of household head	-0.0134***	0.00850	-0.0132***	0.00772	-0.0129***	0.0108*	-0.0106***	0.00543
	(0.00301)	(0.00775)	(0.00290)	(0.00712)	(0.00273)	(0.00643)	(0.00278)	(0.00794)
Number of household members	-0.0542	-0.0674	-0.0505	-0.0726	-0.0485	-0.0855*	-0.0446	-0.0637
	(0.0396)	(0.0506)	(0.0382)	(0.0481)	(0.0320)	(0.0498)	(0.0329)	(0.0639)
% Sawah Area					1.123	0.381	3.270**	-4.717
					(1.171)	(2.166)	(1.663)	(3.512)
Number farms per Ha					-0.132	0.456	0.430	0.0108
					(0.497)	(0.675)	(0.813)	(1.815)
% Sawah Area* Farms/Ha					0.646	-0.778	-0.713	1.874
					(0.840)	(1.488)	(1.483)	(3.585)
Constant	10.63***	0.476	10.46***	0.0563	10.89***	-1.726	9.871***	2.014
	(0.724)	(2.482)	(0.774)	(1.674)	(0.645)	(1.502)	(0.828)	(1.744)
Province FE	yes	yes	yes	yes	yes	yes	yes	yes
Interaction with Islands Dummy	no	no	yes	yes	no	no	yes	yes
Observations	2,505	2,505	2,505	2,505	2,598	2,598	2,598	2,598
Wald chi2	163.33		167.38		430.7		251.95	

Standard errors clustered at district level

See data appendix for details on the samples. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83; Sawah Area = % of agricultural land available for production of wet paddy (rice) in 1963; Farms/Ha = Total number of farms in 1963/ Total Agricultural Land (Ha) in 1963.

Table 20: Correlation between District Characteristics in 1971 and Treatment Intensity in 1982

District Characteristics from 1971 and Treatment Intensity

	Area Bimas/Inmas			Alternative Treatment Variable		
	(1)	(2)	(3)	(4)	(5)	(6)
Total Population 1961	1.76e-07**	1.73e-07**	1.18e-07	1.17e-07	9.54e-08	4.32e-08
	(7.30e-08)	(7.52e-08)	(7.70e-08)	(7.22e-08)	(7.36e-08)	(7.56e-08)
% Male Population in Labor Force		0.139	0.208		-0.0801	-0.0125
		(0.159)	(0.159)		(0.155)	(0.156)
% Female Population in Labor Force		-0.0244	-0.0142		0.00720	0.0187
		(0.0985)	(0.0970)		(0.0964)	(0.0952)
% Male Population in Salary Employment		-0.0311	-0.0631		0.154	0.127
		(0.141)	(0.139)		(0.138)	(0.136)
% Female Population in Salary Employment		0.0258	0.0521		0.0703	0.0977
		(0.125)	(0.123)		(0.122)	(0.121)
% Population born in different Province		0.492	0.464		-0.290	-0.327
		(0.376)	(0.370)		(0.368)	(0.363)
% Population living in Urban Area		-0.240	-0.204		0.122	0.168
		(0.243)	(0.241)		(0.238)	(0.236)
School Construction Program Intensity			-0.0196			-0.00813
			(0.0201)			(0.0197)
W&S Supply Intensity			-0.159*			-0.178*
			(0.0930)			(0.0913)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Test of Joint Significance		**0.79**	**1.35**		**1.27**	**1.66**
P-value		**0.5816**	**0.2252**		**0.275**	**0.1126**
Observations	170	170	170	170	170	170
Adjusted R-squared	**0.651**	**0.645**	**0.657**	**0.449**	**0.453**	**0.468**

Standard errors clustered at the district level

Bimas/Inmas refers to the % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83. Alternative Treatment Variable refers to the % of total agricultural land under any type of Bimas/Inmas intensification in 1982/83. The district level characteristics come from the 1971 census.

Table 21: Means by Cohort and Treatment Intensity (4 quartiles of treatment distribution): Placebo Tests and Experiment of Interest

PANEL A. Placebo Tests					
				Differences	
Treatment	*Cohort 33-42*	*Cohort 43-52*	T2 - T1	T3 - T1	T4 - T1
1	0.39	0.29	-0.099	-0.099	-0.099
2	0.49	0.41	-0.082		
3	0.37	0.29		-0.083	
4	0.43	0.33			-0.106
Difference-in-Difference			***0.016***	***0.015***	***-0.007***
Treatment	*Cohort 33-42*	*Cohort 53-62*	T2 - T1	T3 - T1	T4 - T1
1	0.39	0.24	-0.145	-0.145	-0.145
2	0.49	0.37	-0.124		
3	0.37	0.22		-0.154	
4	0.43	0.28			-0.157
Difference-in-Difference			***0.020***	***-0.009***	***-0.012***
PANEL B. Experiment of Interest				**Differences**	
Treatment	*Cohort 33-42*	*Cohort 1963>*	T2 - T1	T3 - T1	T4 - T1
1	0.39	0.29	-0.094	-0.094	-0.094
2	0.49	0.23	-0.260		
3	0.37	0.12		-0.256	
4	0.43	0.16			-0.273
Difference-in-Difference			***-0.166***	***-0.162***	***-0.178***

Table 22: Effect of the Green Revolution on AM instrumenting with Agricultural Characteristics of 1963

Dep. Var: Arranged Marriage

	IV	IV	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
cohort 33-42				0.0618	-0.0156	0.0511
				(0.126)	(0.155)	(0.155)
cohort 43-52	-0.0840	-0.209	-0.223	-0.0270	-0.214	-0.152
	(0.127)	(0.169)	(0.171)	(0.132)	(0.147)	(0.128)
cohort 53-62	-0.0930	-0.270	-0.281	-0.0223	-0.289	-0.234
	(0.179)	(0.230)	(0.229)	(0.190)	(0.216)	(0.181)
cohort 63>	0.0228	-0.392	-0.402	0.102	-0.383*	-0.320
	(0.190)	(0.247)	(0.249)	(0.194)	(0.225)	(0.203)
cohort 33-42*Area BIMAS/INMAS				-0.246*	-0.200	-0.174
				(0.137)	(0.148)	(0.159)
cohort 43-52*Area BIMAS/INMAS	0.0253	0.121	0.184	-0.213	-0.0886	-0.0138
	(0.163)	(0.181)	(0.196)	(0.166)	(0.152)	(0.146)
cohort 53-62*Area BIMAS/INMAS	**-0.0750**	**0.0594**	**0.158**	**-0.325**	**-0.139**	**-0.0143**
	(0.223)	(0.246)	(0.270)	(0.240)	(0.231)	(0.226)
cohort 63>*Area BIMAS/INMAS	**-0.358**	**-0.0523**	**0.0246**	**-0.614****	**-0.275**	**-0.186**
	(0.240)	(0.271)	(0.289)	(0.246)	(0.247)	(0.248)
Female	0.148***	0.148***	0.148***	0.151***	0.152***	0.152***
	(0.0123)	(0.0123)	(0.0123)	(0.0124)	(0.0124)	(0.0124)
Constant	-0.274	-34.09***	0.417	-0.332**	-48.71***	0.0256
	(0.175)	(5.719)	(0.271)	(0.159)	(10.97)	(0.172)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Int. w/nin	No	Yes	Yes	No	Yes	Yes
Int. w/wssp	No	No	Yes	No	No	Yes
Observations	7883	7883	7883	9068	9068	9068
R-squared	0.263	0.265	0.266	0.278	0.280	0.281
F-statistics of excluded instruments						
cohort 33-42*Area BIMAS/INMAS				215.22	165.82	153.01
cohort 43-52*Area BIMAS/INMAS	282.99	231.65	215.2	251.78	207.14	193.04
cohort 53-62*Area BIMAS/INMAS	290.62	249.39	230.5	251.68	216.39	199.87
cohort 63>*Area BIMAS/INMAS	270.57	227.02	211.76	233.85	195.54	182.29

Standard errors clustered at the district level

For columns 1 to 4, the omitted cohort is the cohort born 1933-43. For columns 5 to 8, the omitted cohort is the cohort born before 1933. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83. Excluded instruments used in columns 3 and 4: Sawah Area = % of agricultural land available for production of wet paddy (rice) in 1963; Farms/Ha = Total number of farms in 1963/ Total Agricultural Land (Ha) in 1963.

Table 23: Effect of the Green Revolution on AM using splines

Dep. Var: Arranged Marriage								
	Treatment = Area Bimas/Inmas				Treatment = Alternative Measure			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
cohort 33-42			-0.0594	-0.293***			-0.0671	-0.283**
			(0.0573)	(0.110)			(0.0711)	(0.134)
cohort 43-52	-0.0704	-0.0758	-0.132***	-0.358***	-0.0289	-0.0450	-0.0988	-0.315**
	(0.0505)	(0.115)	(0.0403)	(0.125)	(0.0436)	(0.119)	(0.0645)	(0.157)
cohort 53-62	-0.223**	-0.309*	-0.278***	-0.599***	-0.153***	-0.261*	-0.214***	-0.535***
	(0.0869)	(0.173)	(0.0679)	(0.180)	(0.0581)	(0.156)	(0.0795)	(0.188)
cohort 63>	-0.155*	-0.252	-0.207**	-0.541***	-0.130**	-0.209	-0.191*	-0.489**
	(0.0817)	(0.189)	(0.0919)	(0.189)	(0.0591)	(0.203)	(0.0983)	(0.235)
cohort 33-42*Treatment 2			-0.0366	-0.0632			-0.0681	-0.0365
			(0.0788)	(0.0693)			(0.0795)	(0.0794)
cohort 33-42*Treatment 3			-0.0906	-0.0973			-0.0823	-0.0545
			(0.0623)	(0.0680)			(0.0698)	(0.0733)
cohort 33-42*Treatment 4			-0.105*	-0.0983			-0.0779	-0.0500
			(0.0601)	(0.0605)			(0.0759)	(0.0758)
cohort 43-52*Treatment 2	-0.00422	0.0157	-0.0377	-0.0467	-0.0802	-0.0793	-0.142*	-0.114
	(0.0581)	(0.0700)	(0.0580)	(0.0645)	(0.0513)	(0.0553)	(0.0740)	(0.0715)
cohort 43-52*Treatment 3	0.0189	0.0533	-0.0675	-0.0396	-0.0107	-0.000131	-0.0862	-0.0510
	(0.0541)	(0.0708)	(0.0492)	(0.0647)	(0.0466)	(0.0555)	(0.0658)	(0.0753)
cohort 43-52*Treatment 4	-0.0118	0.0193	-0.108*	-0.0729	-0.0157	-0.00164	-0.0850	-0.0462
	(0.0557)	(0.0646)	(0.0559)	(0.0600)	(0.0534)	(0.0576)	(0.0790)	(0.0818)
cohort 53-62*Treatment 2	0.0722	0.0709	0.0358	0.00685	-0.0254	-0.0250	-0.0947	-0.0672
	(0.0859)	(0.103)	(0.0835)	(0.100)	(0.0687)	(0.0815)	(0.0931)	(0.0904)
cohort 53-62*Treatment 3	0.0836	0.0792	-0.00641	-0.0202	0.0197	0.0228	-0.0641	-0.0363
	(0.0903)	(0.109)	(0.0755)	(0.0961)	(0.0600)	(0.0833)	(0.0811)	(0.0893)
cohort 53-62*Treatment 4	0.0658	0.0748	-0.0371	-0.0250	0.0131	0.0178	-0.0654	-0.0359
	(0.0932)	(0.110)	(0.0819)	(0.101)	(0.0699)	(0.0863)	(0.0987)	(0.0984)
cohort 63>*Treatment 2	**-0.115**	**-0.106**	**-0.156**	**-0.179***	**-0.158****	**-0.163****	**-0.222****	**-0.204***
	(0.0843)	(0.0999)	(0.0988)	(0.104)	(0.0703)	(0.0817)	(0.104)	(0.106)
cohort 63>*Treatment 3	**-0.106**	**-0.0740**	**-0.196***	**-0.178**	**-0.152****	**-0.114**	**-0.234****	**-0.175**
	(0.0837)	(0.106)	(0.0996)	(0.110)	(0.0622)	(0.0847)	(0.102)	(0.114)
cohort 63>*Treatment 4	**-0.146**	**-0.0952**	**-0.245****	**-0.197***	**-0.145****	**-0.101**	**-0.218***	**-0.155**
	(0.0899)	(0.108)	(0.105)	(0.113)	(0.0722)	(0.0868)	(0.115)	(0.120)
Female	0.149***	0.149***	0.151***	0.153***	0.148***	0.148***	0.151***	0.152***
	(0.0126)	(0.0127)	(0.0127)	(0.0126)	(0.0125)	(0.0126)	(0.0125)	(0.0125)
Constant	0.152***	-0.365***	0.286***	-0.0754	-0.163***	-0.262***	0.151***	-0.0793
	(0.0561)	(0.103)	(0.0678)	(0.109)	(0.0460)	(0.0943)	(0.0466)	(0.124)
District FE	yes	yes	yes	yes	yes	yes	yes	yes
District Controls	no	yes	no	yes	no	yes	no	yes
School Const. Program	no	yes	no	yes	no	yes	no	yes
Water & Sanitation Supply	no	yes	no	yes	no	yes	no	yes
Observations	7,883	7,883	9,068	9,068	7,883	7,883	9,068	9,068
R-squared	0.264	0.268	0.278	0.285	0.264	0.268	0.278	0.285

Standard errors clustered at the district level

For columns 1 to 4, the omitted cohort is the cohort born 1933-43. For columns 5 to 8, the omitted cohort is the cohort born before 1933. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83. The alternative treatment measure refers to the % of total agricultural land covered by any type of intensification.

Table 24: Effect of the Green Revolution on AM allowing for non-linear (quadratic) effects

Dep. Var: Arranged Marriage								
	Treatment = Area Bimas/Inmas				Treatment = Alternative Measure			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
cohort 33-42			-0.0777	-0.282*			-0.124*	-0.330***
			(0.0966)	(0.148)			(0.0644)	(0.113)
cohort 43-52	-0.104	-0.142	-0.182**	-0.410**	-0.0805	-0.114	-0.204***	-0.432***
	(0.0914)	(0.142)	(0.0881)	(0.161)	(0.0501)	(0.111)	(0.0705)	(0.143)
cohort 53-62	-0.345***	-0.434**	-0.414***	-0.711***	-0.168**	-0.269*	-0.287***	-0.597***
	(0.123)	(0.211)	(0.135)	(0.235)	(0.0771)	(0.139)	(0.0971)	(0.179)
cohort 63>	-0.137	-0.231	-0.195	-0.497**	-0.155*	-0.263	-0.270***	-0.595***
	(0.132)	(0.242)	(0.157)	(0.240)	(0.0818)	(0.181)	(0.103)	(0.219)
cohort 33-42*Treatment			0.0125	-0.161			-0.00541	0.0336
			(0.442)	(0.388)			(0.290)	(0.292)
cohort 43-52*Treatment	0.214	0.241	0.213	0.0693	-0.0375	0.0350	-0.0289	0.0809
	(0.353)	(0.388)	(0.384)	(0.376)	(0.235)	(0.233)	(0.318)	(0.254)
cohort 53-62*Treatment	**0.682***	**0.634**	**0.672**	**0.460**	**-0.0301**	**-0.0323**	**-0.0304**	**0.00707**
	(0.388)	(0.444)	(0.502)	(0.486)	(0.367)	(0.349)	(0.461)	(0.392)
cohort 63>*Treatment	**-0.0848**	**-0.243**	**-0.150**	**-0.473**	**-0.771****	**-0.587**	**-0.771**	**-0.548**
	(0.470)	(0.520)	(0.552)	(0.526)	(0.376)	(0.364)	(0.504)	(0.446)
cohort 33-42*Treatment^2			-0.0947	0.0665			-0.0936	-0.133
			(0.377)	(0.333)			(0.377)	(0.341)
cohort 43-52*Treatment^2	-0.191	-0.182	-0.267	-0.0991	0.215	0.140	0.115	0.00277
	(0.291)	(0.308)	(0.338)	(0.302)	(0.270)	(0.278)	(0.381)	(0.297)
cohort 53-62*Treatment^2	**-0.510***	**-0.451**	**-0.583**	**-0.375**	**0.195**	**0.179**	**0.0890**	**0.0371**
	(0.308)	(0.329)	(0.429)	(0.388)	(0.375)	(0.378)	(0.495)	(0.412)
cohort 63>*Treatment^2	**-0.0850**	**0.0982**	**-0.108**	**0.220**	**0.985****	**0.827****	**0.883**	**0.684**
	(0.391)	(0.405)	(0.459)	(0.405)	(0.402)	(0.394)	(0.574)	(0.485)
Female	0.148***	0.148***	0.151***	0.153***	0.148***	0.148***	0.150***	0.152***
	(0.0125)	(0.0126)	(0.0126)	(0.0125)	(0.0124)	(0.0125)	(0.0125)	(0.0125)
Constant	-0.00575	-0.343***	0.0478	-0.241**	0.0876	-0.324***	0.202**	0.135
	(0.135)	(0.0898)	(0.161)	(0.105)	(0.0624)	(0.0876)	(0.0806)	(0.170)
District FE	yes	yes	yes	yes	yes	yes	yes	yes
District Controls	no	yes	no	yes	no	yes	no	yes
School Const. Program	no	yes	no	yes	no	yes	no	yes
Water & Sanitation Supply	no	yes	no	yes	no	yes	no	yes
Observations	7,883	7,883	9,068	9,068	7,883	7,883	9,068	9,068
R-squared	0.265	0.268	0.279	0.285	0.265	0.268	0.278	0.285
Total Effect 33-42			-0.121	-0.0673			-0.0625	-0.0475
			(0.121)	(0.112)			(0.105)	(0.121)
Total Effect 43-52	-0.0549	-0.0461	-0.163	-0.0703	0.0939	0.121	0.0413	0.0826
	(0.0887)	(0.0913)	(0.122)	(0.0968)	(0.0943)	(0.0936)	(0.129)	(0.115)
Total Effect 53-62	-0.0363	-0.0608	-0.150	-0.0691	0.0886	0.0769	0.0239	0.0297
	(0.108)	(0.113)	(0.158)	(0.144)	(0.153)	(0.139)	(0.193)	(0.168)
Total Effect 63>	**-0.205**	**-0.188**	**-0.302***	**-0.163**	**-0.171**	**-0.0827**	**-0.233**	**-0.131**
	(0.135)	(0.145)	(0.166)	(0.146)	(0.156)	(0.150)	(0.199)	(0.183)

Standard errors clustered at the district level

For columns 1 to 4, the omitted cohort is the cohort born 1933-43. For columns 5 to 8, the omitted cohort is the cohort born before 1933. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83. The alternative treatment measure refers to the % of total agricultural land covered by any type of intensification. The total effect is calculated at the mean of the treatment variable.

Table 25: Effect of the Green Revolution on Education using Census 1995

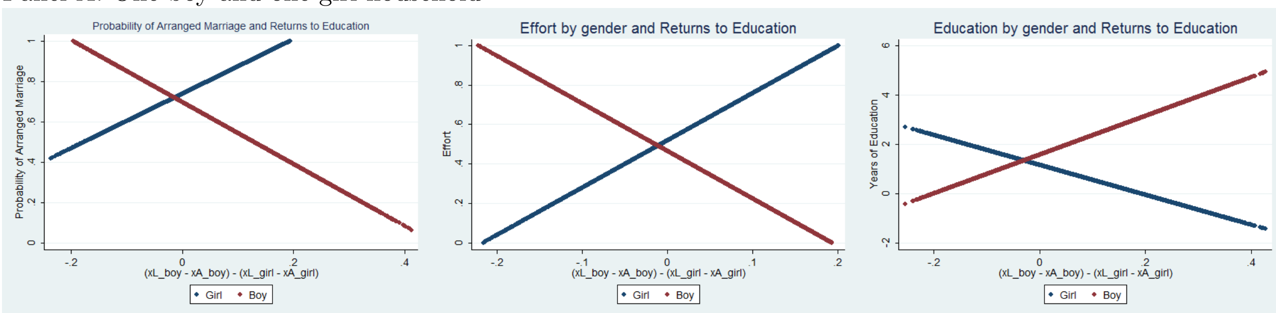
Dep. Var: Education	OLS	OLS	IV	IV	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(9)
cohort 33-42					1.255***	1.759***	1.353***	1.799***
					(0.152)	(0.224)	(0.315)	(0.438)
cohort 43-52	1.220***	1.734***	1.165***	1.663***	2.478***	3.552***	2.524***	3.508***
	(0.141)	(0.167)	(0.248)	(0.315)	(0.187)	(0.308)	(0.397)	(0.558)
cohort 53-62	2.346***	2.593***	2.168***	2.292***	3.605***	4.477***	3.546***	4.334***
	(0.161)	(0.229)	(0.314)	(0.448)	(0.189)	(0.307)	(0.419)	(0.636)
cohort 63>	3.746***	3.593***	3.166***	2.663***	5.010***	5.428***	4.753***	4.788***
	(0.210)	(0.258)	(0.445)	(0.665)	(0.217)	(0.314)	(0.468)	(0.726)
cohort 33-42*Area BIMAS/INMAS					0.350*	0.0914	0.327	0.0323
					(0.178)	(0.193)	(0.483)	(0.508)
cohort 43-52*Area BIMAS/INMAS	0.540***	0.350**	0.609*	0.422	0.884***	0.407	0.982	0.448
	(0.158)	(0.155)	(0.349)	(0.358)	(0.235)	(0.256)	(0.654)	(0.673)
cohort 53-62*Area BIMAS/INMAS	0.431**	0.331*	0.671	0.668	0.768***	0.374	1.005	0.541
	(0.195)	(0.200)	(0.461)	(0.502)	(0.248)	(0.262)	(0.705)	(0.758)
cohort 63>*Area BIMAS/INMAS	**0.877*** **	**0.921*** **	**1.743*** **	**2.028*** **	**1.209*** **	**0.963*** **	**1.874** **	**1.823** **
	(0.260)	(0.261)	(0.657)	(0.743)	(0.283)	(0.290)	(0.788)	(0.872)
Female	-1.374***	-1.375***	-1.374***	-1.374***	-1.419***	-1.419***	-1.419***	-1.419***
	(0.0335)	(0.0335)	(0.0334)	(0.0335)	(0.0316)	(0.0317)	(0.0317)	(0.0318)
Constant	4.787***	5.144***	21.02***	12.28***	3.351***	3.877***	3.788***	3.711***
	(0.160)	(0.196)	(1.617)	(0.534)	(0.140)	(0.266)	(0.355)	(0.360)
District of birth FE	yes	yes	yes	yes	yes	yes	yes	yes
Int. w/other programs	no	yes	no	yes	no	yes	no	yes
Observations	327,404	327,404	327,404	327,404	360,683	360,683	360,683	360,683
R-squared	0.218	0.219	0.218	0.218	0.284	0.284	0.284	0.284
F-statistics of excluded instruments								
cohort 33-42*Area BIMAS/INMAS							7423.58	6576.85
cohort 43-52*Area BIMAS/INMAS			9606.39	8389.43			7919.82	6925.27
cohort 53-62*Area BIMAS/INMAS			10037.12	8697.52			8265.09	7171.96
cohort 63>*Area BIMAS/INMAS			9987.86	8527.24			8165.28	6961.17

Standard errors clustered at the district level

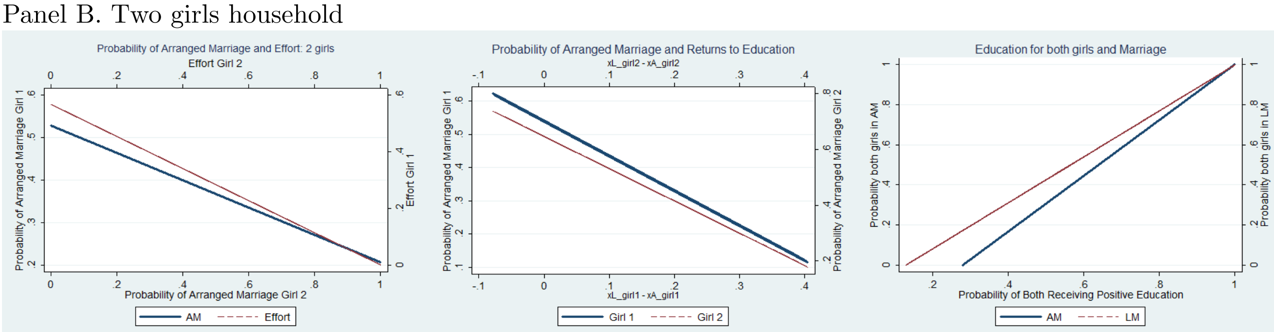
For columns 1 to 4, the omitted cohort is the cohort born 1933-43. For columns 5 to 8, the omitted cohort is the cohort born before 1933. Bimas/Inmas = % of total agricultural wet paddy (rice) land under any type of Bimas/Inmas intensification in 1982/83. Excluded instruments used in columns 3, 4, 8 and 9: Sawah Area = % of agricultural land available for production of wet paddy (rice) in 1963; Farms/Ha = Total number of farms in 1963/ Total Agricultural Land (Ha) in 1963.

Figure 13: Numerical Comparative Statics: Two children

Panel A. One boy and one girl household



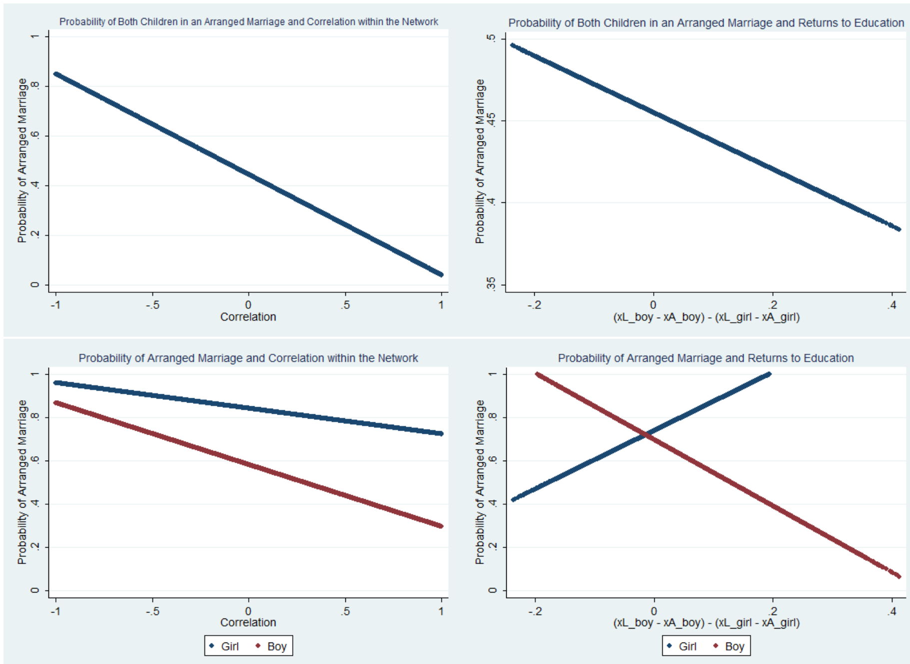
Panel B. Two girls household



One boy and one girl: The parameters used for the simulations are $p_g = 0.1$, $p_b = .105$ $\beta \in [0.8, 1]$, $\varphi = 0.35$, $d \in [0.15, 0.65]$, $d_b = d_g \in [0.01, 0.55]$, $e_{high} = 0.1$, $e_{low} = 0$, $\sigma_\delta^2 \in [0.1, 3]$, $\varrho_{g_j,b_n}(e_{g,j} = 1, e_{b,n} = 1) = 1$, $\varrho_{g_j,b_n}(e_{g,j} = 0, e_{b,n} = 1) = \varrho_{g_j,b_n}(e_{g,j} = 1, e_{b,n} = 0) = \varrho_{g_j,b_n}(e_{g,j} = 0, e_{b,n} = 0) = 0$ $\alpha_{bL} = \alpha_{gL} \in [0.001, 0.25]$, $\alpha_{bA} = \alpha_{gA} \in [0.001, 0.25]$, x_{gL}, x_{bL}, x_{gA} and x_{bA} are bounded such that education is increasing in the returns.

Two girls: The parameters used for the simulations are $p_{g1} = p_{g2} = 0.1$, $\beta \in [0.8, 1]$, $\varphi = 0.35$, $d \in [0.15, 0.65]$, $d_{g1} = d_{g2} \in [0.01, 0.55]$, $e_{high} = 0.1$, $e_{low} = 0$, $\sigma_\delta^2 \in [0.1, 3]$, $correlation = 1$, $\alpha_{bL} = \alpha_{gL} \in [0.001, 0.25]$, $\alpha_{bA} = \alpha_{gA} \in [0.001, 0.25]$, x_{g1L}=x_{g2L}, and $x_{g1A} = x_{g2A}$ are bounded such that education is increasing in the returns.

Figure 14: Numerical Comparative Statics allowing for Correlation between households to vary: 2 children



1 boy and 1 girl: The parameters used for the simulations are $p_g = 0.1$, $p_b = .105$ $\beta \in [0.8, 1]$, $\varphi = 0.35$, $d \in [0.15, 0.65]$, $d_b = d_g \in [0.01, 0.55]$, $e_{high} = 0.1$, $e_{low} = 0$, $\sigma^2_\delta \in [0.1, 3]$, $\varrho_{g_j,b_n}(e_{g,j} = 1, e_{b,n} = 1) \in [-1, 1]$, $\varrho_{g_j,b_n}(e_{g,j} = 0, e_{b,n} = 1) = \varrho_{g_j,b_n}(e_{g,j} = 1, e_{b,n} = 0) = \varrho_{g_j,b_n}(e_{g,j} = 0, e_{b,n} = 0) = 0$, $\alpha_{bL} = \alpha_{gL} \in [0.001, 0.25]$, $\alpha_{bA} = \alpha_{gA} \in [0.001, 0.25]$, x_{gL}, x_{bL}, x_{gA} and x_{bA} are bounded such that education is increasing in the returns.

Table 26: Divorce Rates by Cohort Indonesia

Dep. Var: Divorce	OLS	OLS	OLS
	(1)	(2)	(3)
Female	0.0643***	0.0649***	0.0658***
	(0.00536)	(0.00534)	(0.00533)
Duration	-0.0191***	-0.0192***	-0.0191***
	(0.000503)	(0.000505)	(0.000503)
Cohort 33-42	-0.116***	-0.116***	-0.119***
	(0.0157)	(0.0157)	(0.0159)
Cohort 42-48	-0.221***	-0.221***	-0.223***
	(0.0134)	(0.0134)	(0.0134)
Cohort 49-55	-0.354***	-0.355***	-0.351***
	(0.0142)	(0.0141)	(0.0139)
Cohort 56-62	-0.482***	-0.483***	-0.478***
	(0.0153)	(0.0153)	(0.0153)
Cohort 63-69	-0.624***	-0.625***	-0.626***
	(0.0178)	(0.0177)	(0.0181)
Cohort 70>	-0.719***	-0.722***	-0.734***
	(0.0204)	(0.0206)	(0.0212)
Arranged	-0.0371	0.0425**	0.0458**
	(0.0300)	(0.0177)	(0.0179)
Cohort 33-42*Arranged	0.0179	0.0141	0.0112
	(0.0231)	(0.0231)	(0.0233)
Cohort 42-48*Arranged	0.0463*	0.0419*	0.0378
	(0.0242)	(0.0245)	(0.0242)
Cohort 49-55*Arranged	0.0797***	0.0732***	0.0711***
	(0.0243)	(0.0240)	(0.0244)
Cohort 56-62*Arranged	0.0956***	0.0864***	0.0856***
	(0.0245)	(0.0247)	(0.0247)
Cohort 63-69*Arranged	0.120***	0.116***	0.107***
	(0.0312)	(0.0334)	(0.0327)
Cohort 70>*Arranged	0.0589	0.0529	0.0508
	(0.0390)	(0.0393)	(0.0400)
Constant	0.735***	0.725***	0.706***
	(0.0203)	(0.0219)	(0.0186)
District FE	No	No	Yes
Province FE	Yes	Yes	No
Interaction Prov FE*Arr	Yes	No	No
Observations	12,206	12,206	12,206
R-squared	0.395	0.391	0.413

Standard errors clustered at the district level

The first column presents the results of $D_{ipc} = =\beta_0 + \gamma_c + \eta_p + \beta_1 AM_{ipc} + \sum_c (\gamma_c * AM_{ipc})\beta_{2,c} + \sum_p (\eta_p * AM_{ipc})\beta_{3,p} + \beta_4 female_{ipc} + \beta_5 duration_{ipc} + \varepsilon_{ipc}$. The second column does not interact the province fixed effects with the dummy variable for arranged marriages. And the third column uses district fixed effects instead of province fixed effects.