KU LEUVEN

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES
DPS16.02

MARCH 2016

Where did it go wrong? Marriage and divorce in Malawi

Laurens CHERCHYE, Bram DE ROCK, Selma TELALAGIC WALTHER and Frederic VERMEULEN

Public Economics

Faculty of Economics
and Business

Where did it go wrong? Marriage and divorce in Malawi*

Laurens Cherchye† Bram De Rock‡ Frederic Vermeulen§ Selma Walther¶

July 2018

Abstract

Do individuals marry and divorce for economic reasons? Can we measure the economic attractiveness of a person's marriage market? We answer these questions using a structural model of consumer-producer households that is applied to rich data from Malawi. Using revealed preference conditions for a stable marriage market, we define the economic attractiveness of a potential match as the difference between the potential value of consumption and leisure with the new partner and the value of consumption and leisure in the current marriage. We estimate this marital instability measure for every pair in geographically defined marriage markets in 2010. We find that the marital instability measure is predictive of future divorces, particularly for women with attractive outside options. We further show that this estimated effect on divorce is mitigated by the woman's age, and by a lack of men, relative to women, in the marriage market, showing that these factors interact with the economic attractiveness of the remarriage market in decisions about divorce. These findings provide out-of-sample validation of our model and the first structural evidence that the value of the marriage market matters for divorce decisions.

Keywords: marriage market, divorce, Malawi, agricultural production, revealed preference.

JEL classification: D11, D12, D13, J12.

*We thank Andrew Beauchamp, Ian Crawford, James Fenske, Arthur Lewbel, Costas Meghir, Alessandra Voena and seminar participants at University College London, Oxford, Rice, Texas A&M, IIES Stockholm, Tilburg, Warwick, Essex, Boston College, LISER, the Econometric Society European Meeting of 2016, the North American Meeting of the Econometric Society of 2017, and the European Society for Population Economics of 2018 for useful discussion.

†Department of Economics, University of Leuven (KU Leuven). E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven-kortrijk.be. Laurens Cherchye gratefully acknowledges the European Research Council (ERC) for his Consolidator Grant 614221. Part of this research is also funded by the FWO (Research Fund-Flanders) and the Research Fund of the KU Leuven (BOF).

‡ECARES, Université Libre de Bruxelles and Department of Economics, University of Leuven (KU Leuven). Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges FWO and BELSPO for their financial support.

§Department of Economics, University of Leuven (KU Leuven) and Universidad del Pacifico. Naamsestraat 69, B-3000 Leuven, Belgium. E-mail: frederic.vermeulen@kuleuven.be. Frederic Vermeulen gratefully acknowledges financial support from the Research Fund KU Leuven through the grant STRT/12/001 and from the FWO through the grant G057314N.

¶Department of Economics, University of Warwick, Coventry CV47AL, s.walther@warwick.ac.uk.

1 Introduction

Becker (1973, 1974) convincingly argued that the institution of marriage can be analyzed by means of modern microeconomic theory. In his ground-breaking work, as well as in subsequent work by Becker, Landes and Michael (1977), the concept of the marriage market is introduced, which rests on the simple but powerful assumption that individuals are rational utility maximizers who compete as they seek mates. This framework implies that each individual looks for the best mate subject to the restrictions imposed by the marriage market. An important concept in this theory is gains to marriage, which depend on a given union as well as the opportunities provided by the marriage market as a whole. While companionship and the production of children are important components of marital gains, there are also considerable economic gains to marriage, such as the sharing of public goods and the division of labor within unions (see Browning, Chiappori and Weiss, 2014, for an extensive discussion).

In this paper, we focus on the relationship between the economic gains to marriage, and divorce and remarriage. Individuals seek to find the best match in the marriage market, and better outside options in terms of one's marriage market will affect both intrahousehold sharing in the current match, and subsequent divorce, if opportunities on the marriage market dominate the allocation in the current marriage. Although the idea that outside options on the marriage market affect divorce decisions is widely accepted, convincing evidence is lacking. We provide a structural measure of the value of the remarriage market, and show that it predicts future divorce, in an out-of-sample test of the model.

In estimating our model, we fix our attention on households in Malawi, a context where divorce is common and remarriage is socially acceptable. Lifetime divorce probabilities are between 40% and 65%, and remarriage is almost universal: within two years of divorce, over 40% of women remarry, with this figure reaching almost 90% after ten years (Reniers 2003; see also the discussion in Section 2). Marriages also tend to happen within neighboring villages, which allows the accurate definition of marriage markets.

Marriage market. In recent work, Cherchye, Demuynck, De Rock and Vermeulen (2017a) analyze the impact of the marriage market on the intrahousehold distribution of resources, focusing on the gains from public goods. These authors combine the static collective model of household consumption (Chiappori 1988, 1992) with the assumption of a stable marriage market, the latter relating directly to the ideas in Becker (1973, 1974) and Becker, Landes and Michael (1977). The model quantifies the outside options of a myopic spouse and subsequently relates this to that spouse's share of household resources. These outside options improve with one's productivity, which implies that the marriage market can explain the widely observed positive relationship between wages and the share of household resources consumed (see, for example, Blundell, Chiappori, Magnac and Meghir, 2007, Cherchye, De Rock and Vermeulen, 2012, and Cherchye, De Rock, Lewbel and Vermeulen, 2015).

Dunbar, Lewbel and Pendakur (2013) do not reject the implications of the collective model among Malawi households, which allows them to identify the intrahousehold sharing of resources. However, they do not explicitly model the marriage market. To do this, we take as our starting point the framework of Cherchye, Demuynck, De Rock and Vermeulen (2017a). More precisely, given the importance of agricultural production in Malawi (see Section 2 for more details), we present a structural model of consumer-producer households that integrates economic gains to marriage, both in terms of public goods and the division of labor in household production, and that accounts for the intrahousehold allocation of resources in the context of a marriage market (see also Apps and Rees, 1996, and Chiappori, 1997). As we explain below, our model yields precise measures of the value of an individual's marriage market.

At this point we want to note that, although our model is static and assumes stability on the marriage market, this is not necessarily a contradiction with the widespread observation that households divorce. It simply implies that myopic individuals do not take into account future shocks that may change their current (and future) choices. These static models are popular in the literature (see for instance Browning, Chiappori and Weiss, 2014) and can be considered as a building block for more advanced models that focus on the intertemporal aspects of household decisions (as in, e.g., Mazzocco, 2007, who focuses on intertemporal consumption choices). Also, it is important to note that our model performs remarkably well empirically: there is no *a priori* reason to expect the predictive power on future divorces that we find.

Predicting divorce. Our model yields two structural measures of the value of an individual's outside option, which we term *marital instability indices*: the first index captures how much better off (in consumption terms) the individual would be if single (the Individual Rationality (*IR)* index), while the second index measures how much better off the individual would be if he/she remarried another individual in the same marriage market (the Blocking Pair (*BP)* index). Computing the *BP* index for each possible pair within each marriage market, we then take the maximum of an individual's set of *BP* indices to obtain an estimate of the economic value of the (re)marriage market that reflects the individual's most attractive remarriage option.[1] We estimate these instability indices for each married

[1] We also consider the average of an individual's *BP* indices, and the 95th percentile of an individual's *BP* indices, with very similar results - see Section 6.

individual in the first wave of our data (2010). Using the 2013 wave of the dataset, we are also able to observe if the individual divorces in the next three years. If the individuals have divorced by 2013, this indicates that they experienced an important shock between 2010 and 2013; for example, a change in the economic opportunities on the marriage market, or a deterioration of match quality in the marriage. We link our measures of instability to these observed subsequent divorces, which sheds light on the relationship between economic gains to marriage and divorce, and is also an out-of-sample test of the validity of our structural model.

We find that the wife's BP index significantly predicts subsequent divorce. In particular, a 1 percentage point increase in the wife's BP index, as a proportion of her household income, raises the probability of divorce by 1.4 percentage points on average. This is an economically significant effect, as the per-year divorce probability is 8.5%.[2] Interestingly, this significant association cannot be explained by spouses' wages, land income or nonlabor income which, alongside intrahousehold sharing, are the key determinants of the BP index in the structural model. This suggests that intrahousehold sharing plays an important role in the gains to marriage and divorce. As an extension to these results, we also estimate a model that allows the instability indices to have a different effect on divorcing and remaining single, and divorcing and remarrying. Crucially, we find that the wife's BP index is significantly associated with the wife divorcing and remarrying, but not divorcing and remaining single. This is consistent with the intuition that the BP index captures the attractiveness of options on the remarriage market. Therefore, we find that a model-based measure of individuals' outside options on the marriage market correlates with out-of-sample realizations of divorce.

Relation to the literature. Our paper makes two key contributions to the literature. First, from a methodological point of view, it significantly extends the theoretical model in Cherchye, Demuynck, De Rock and Vermeulen (2017a) by also accounting for the economic gains of production decisions in modeling households' behavior. This is particularly relevant for consumer-producer households in developing countries, for which agricultural production activities are prevalent (see, for example, Udry, 1996, Walther, 2018, Apps and Rees, 1996, Chiappori, 1997, and Karlan, Osei, Osei-Akoto and Udry, 2014). A distinguishing feature of our approach is that it belongs to a revealed preference tradition that is free of any parametric assumptions, and optimally allows for heterogeneity in preferences and production technologies. See Samuelson (1938), Afriat (1967), Diewert (1973) and Varian (1982) for early contributions on the revealed preference analysis of household consumption

[2] Modeling divorce as a simple Markov process, and using the proportions of individuals currently married and divorced in the dataset and the remarriage probabilities in Reniers (2003), implies an annual divorce probability of 8.5%.

behavior, and Afriat (1972) and Varian (1984) for the analysis of production behavior. More recently, Cherchye, De Rock and Vermeulen (2007, 2009, 2011) have extended this seminal work towards the analysis of households in the framework of a collective model. Finally, and importantly from a methodological point of view, our revealed preference methods allow us to estimate shadow wages and land prices, which are often missing or suffer from measurement error in empirical applications. As such, we obtain an empirically tractable model that can be applied to a context with consumer-producer households to study household choices and the role of the marriage market.

Second, our empirical application contributes a unique perspective to the active literature on the economic drivers of divorce. Many studies focus on the role of shocks in a reduced form approach. For example, unemployment (Charles and Stephens, 2004, Doiron and Mendolia, 2011, and Eliason, 2012), shocks to earnings capacity (Weiss and Willis, 1997), television access (Chong and La Ferrara, 2009), changes in house prices (Farnham, Schmidt and Sevak, 2011), and lottery winnings (Hankins and Hoekstra, 2011), to name a few, have all been shown to correlate significantly with subsequent divorce, while structural models make more precise the mechanisms behind divorce decisions, but have tended to focus on learning about match quality, or the role of policy changes (Brien, Lillard and Stern, 2006, Jacquemet and Robin, 2012, Bruze, Svarer and Weiss, 2015, and Voena, 2015). Our paper breaks now ground in the literature on divorce in two important ways: first, we combine a structural and reduced form approach, using a structural model to provide a theoretical underpinning to the value of an individual's outside option on the marriage market, and a reduced form approach to correlate this measure with subsequent divorces; second, we provide the first empirical evidence of the well-known intuition that outside options, and in particular the value of one's remarriage market, matter for divorce decisions. More generally, we model the consumption, labor supply and marital status choices of households in a low-income country in a unified way.

Our findings on the role of outside opportunities in triggering divorce complement models where match quality plays an important role in marriage and divorce decisions: individuals can be thought of as matching primarily for economic reasons, but high match quality can compensate for economic "mismatch". However, when match quality erodes, individuals search for a better economic match, and so divorce when there are more economically attractive individuals available in their marriage market (see, e.g., Chiappori, Radchenko and Salanie, 2018, who use economic and non-economic measures of match quality to predict divorce). In fact, we find that match quality matters in addition to our measures of economic gains on the marriage market. For example, we find that the estimated effect of the value of the remarriage market on divorce is mitigated when spouses are older, and is reduced when spouses are assortatively matched on age. The latter result relates to findings on the importance of assortative matching in marriage (see, e.g., Hitsch, Hortacsu and Ariely, 2010, Greenwood, Guner, Kocharkov and Santos, 2014, and Chiappori, Oreffice and Quintana-Domeque, 2017). Finally, and consistent with the literature, the sex ratio is an important determinant of outside options: the relative attractiveness of women's remarriage opportunities turns out to be less predictive of divorce when there are fewer men relative to women in her marriage market (Chiappori, Fortin and Lacroix, 2012).

Structure. The rest of this paper unfolds as follows. Section 2 describes the context of Malawi, which motivates the structure of our model and the empirical analysis. Section 3 introduces our revealed preference methodology for analyzing the stability of marriage. Here, we also define our *IR* and *BP* indices for marriage stability. In Section 4 we discuss the dataset and explain how we construct marriage markets. Section 5 presents summary statistics of the main outcomes of our structural model. These results motivate our key empirical analysis in Section 6, in which we focus on the empirical relationship between the economic gains to matches (captured by our structural *IR* and *BP* indices) and divorce and remarriage probabilities. Section 7 concludes.

2 Malawian context

Malawi is a poor country in Sub-Saharan Africa, with a GDP per capita of $226 in 2013 (World Bank). It ranks 174th out of 187 countries on the 2014 Human Development Index, with an average life expectancy of 55.3 years. The proportion of females with secondary school education is low, at 10.4%. Households in Malawi primarily engage in subsistence agricultural production, with smallholder plots in the region of 0.2-3 hectares (Bignami-Van Assche et al., 2011, Ellis, Kutengule and Nyasulu, 2003). Land is largely passed on through inheritance, often at the time of marriage, and determined by descent, which can be matrilineal or patrilineal (Walther, 2018). The predominant crop grown is maize, and agricultural production involves the joint labor supply of husbands and wives (see Walther, 2017, for more information on labor supply). Individuals' primary assets, and thus determinants of outside options, are their landholdings and capacity for labor supply. These features make it important to take account of households' agricultural production when considering their decision-making.

There are two key reasons why we choose this context to examine the role of economic factors in divorce. First, Malawi is characterized by high divorce rates. Marriage is almost universal (Reniers, 2003), with over 99% of women and 97% of men having married at least once by the age of 30 (Demographic Health Survey Report, 2004). Early marriage is common, with the median age of first marriage at 18 for women and 23 for men (DHS Report, 2004). However, marriage is also unstable, with almost half of all marriages ending within twenty years, a figure much higher than in other African countries, and similar to present-day figures for the U.S. (Reniers, 2003). In this sense, Malawi is characterized by a high turnover of marriages and divorces. One driver of the high divorce rate is that divorce is easy to obtain: spouses seeking divorce need only state that there is no love remaining in the marriage (Mwambene, 2005). An important observation when applying our model is that remarriage is also common, with 40% of women remarrying within two years. Thus, Malawi is characterized by an ease of moving between marriage and divorce, which is consistent with the assumptions of our model presented in Section 3, with no frictions on the marriage market and where outside options are determined by utility on divorce.

Second, marriage is local. Approximately 45% of married individuals are from the village they live in, while a further 25% are from another village within the same district (Malawi IHS 2010). This allows us to be precise about defining the marriage markets within which divorced individuals can look for potential remarriage partners. In particular, we use geographic information about households to construct marriage markets - we discuss this in detail in Section 4.

To get a sense of the reasons for divorce given by individuals in Malawi, Table 1 shows responses given by men and women in the 2008 wave of the Malawi Longitudinal Study of Families and Health (University of Pennsylvania) to the question: What was the main reason your marriage ended? Respondents gave a complete history of their marriages and divorces. The modal response across both men and women is "lack of love". However, unfaithfulness of the spouse is the next most common reason, and is also closely related to the answer "Spouse married someone else", as both involve the presence of an alternative partner. Combining these two categories implies that among men, approximately 43% of divorces occurred due to the presence of another partner, while among women, this figure is 42%.

Table 1: Reasons for divorce, responses in 2008 Malawi Longitudinal Survey of Families and Health

	Men	Women
Lack of love	28.7%	31.5%
Spouse unfaithful	35.8%	21.9%
Spouse did not provide	4.8%	9.6%
Spouse married someone else	7.5%	20.2%
Respondent unfaithful	5.9%	3.6%
Suspected spouse of having HIV	0%	0.2%
Other	17.3%	13.0%
# Recorded divorces	734	977
# Recorded marriages	2566	3186
% Marriages ending in divorce	28.6	26.2

3 Consumption, production and marriage stability

Our method for measuring the instability of marriage takes as a starting point the model of Cherchye, Demuynck, De Rock and Vermeulen (2017a). These authors define a revealed preference characterization of household consumption under stable marriage to analyze the intrahousehold allocation of resources and the gains to marriage in terms of public goods. A novel feature of our analysis is that we integrate household production in this revealed preference framework, thus linking productivity to marriage decisions.

As explained in Section 2, agricultural production is an important dimension of household decisions in developing countries, and Malawi in particular. It is the primary source of livelihood and a crucial determinant of outside options. Moreover, our structural modeling of household production allows us to use shadow wages and land prices in our analysis of marriage stability. This is particularly important in view of our aim to accurately estimate the value of the individuals on the marriage market. The majority of households in Malawi do not perform market work, which means that observed market wages are likely to be upward biased relative to the distribution of wage offers, and will upward bias the estimated productivity of individuals and their value on the marriage market. Our method circumvents this issue by estimating each individuals' productivity on the land, which is a more accurate measure of economic attractiveness for farming households. This also indicates the usefulness of our model for other settings where individuals' productivity on the land is an important factor.

3.1 Notation and components of the structural model

We focus on the marriage stability of couples that consist of a female a and a male b. In what follows, we will often refer to individual $i = a, b$. Let A be a finite set of females and B a finite set of males. The marriage market is defined by a matching function $\sigma : A \cup B \rightarrow A \cup B$. This function satisfies, for all $a \in A$ and $b \in B$,

$$\sigma(a) \in B, \sigma(b) \in A,$$
$$\sigma(a) = b \in B \text{ if and only if } \sigma(b) = a \in A.$$

In words, the function σ assigns to every female or male a partner of the other gender (i.e. $\sigma(a) = b$ and $\sigma(b) = a$). For simplicity we will assume in this methodological section that $|A| = |B|$, which means that all individuals are matched. Actually, it is relatively straightforward to formally include singles in the models below.[3] However, unless there is a shortage on one side of the marriage market, rationalizing the behavior of singles requires an explicit model for frictions on the marriage market, or marriage costs. To focus our discussion, we abstract from these extensions in the theoretical framework, but we allow potential matches to be formed between married individuals and singles in the marriage market in the empirical analysis, so that empirically we allow for the possibility that $|A| \neq |B|$.

Each individual i is assumed to spend his or her total time endowment (denoted by $T^i \in \mathbb{R}_+$) on leisure ($l^i \in \mathbb{R}_+$), market work ($m^i \in \mathbb{R}_+$) and agricultural work on the household's land (denoted by $h^i \in \mathbb{R}_+$).[4] The individual's budget constraint for time is

$$T^i = m^i + h^i + l^i.$$

The price of time is individual i's wage, which we represent by $w^i \in \mathbb{R}_{++}$.

To model agricultural production, we assume that there are three types of inputs: the individuals' time spent on agricultural labor (h^a and h^b), land ($L \in \mathbb{R}_+$) and other inputs ($x \in \mathbb{R}_+$; for example, fertilizer). To take our Malawi context into account, we distinguish between land belonging to the female ($L^a \in \mathbb{R}_+$), land belonging to the male ($L^b \in \mathbb{R}_+$) and joint "household" land ($L^{(a,b)} \in \mathbb{R}_+$):

$$L = L^a + L^b + L^{(a,b)}.$$

[3] Specifically, some of the variables in Propositions 1 and 2 (individual quantities, share of nonlabor income and shadow wages) must be set equal to zero in the case of singles. But the basic structure of the rationalizability conditions in the propositions remains intact.

[4] In the empicial application, we assume that everyone has the same time endowment: $T^i = T$.

The first two types of land are assignable in the post-divorce allocation, while this typically is not the case for jointly owned land. For a given match (a, b), we assume a common price for the three land types, so that the price of L^a, L^b and $L^{(a,b)}$ is given by $z^{(a,b)} \in \mathbb{R}_{++}$. The other input x is assumed to be a Hicksian aggregate with a price that is normalized to unity. The inputs are transformed into an output $y \in \mathbb{R}_+$ by means of an agricultural production function $F\left(h^a, h^b, L, x\right)$. We assume that this function is increasing in its arguments and characterized by constant returns to scale (in line with Pollak and Wachter, 1975). The output associated with agricultural production is again a Hicksian aggregate, with a price that is normalized to unity. Note that we make the assumption that agricultural production is marketable. As such, it is associated with an exogenous normalized price (see also Chiappori, 1997). The household is further associated with nonlabor income $n^{(a,b)} \in \mathbb{R}_+$.

The total income of a household consists of income from market work, agricultural production and nonlabor income. It is allocated to a Hicksian aggregate good with a price that is normalized to unity. This Hicksian aggregate is used for the private consumption of both spouses (denoted by $q_a, q_b \in \mathbb{R}_+$) and the household's consumption of a public good (denoted by $Q \in \mathbb{R}_+$). Examples of private goods are food and clothing, while an example of a public good is expenditure on children. Importantly, the household's consumption of the private good, for example food, equals the sum of the food bought at the market and food produced at home when the household produces less than it consumes, and equals a share of the home produced food when the household produces more than it consumes. Further, by including public consumption, our model effectively captures economies of scale in consumption, which form a prime economic motivation for marriage (in addition to household (agricultural) production).

Finally, each individual i is assumed to derive utility from leisure, private consumption as well as public consumption. The preferences of individual i are represented by a utility function $U^i\left(l^i, q_i, Q\right)$ that is assumed to be continuous, concave and strictly increasing in leisure l^i and private consumption q_i, and increasing in public consumption Q.

3.2 Marriage stability: theoretical characterization

We now define a stable marriage allocation. We say that an allocation is stable if it satisfies three equilibrium conditions.

First, at the production level, we follow the set-up of Chiappori (1997) and assume that each household $(a, \sigma(a))$ is a profit maximizer. This implies that the chosen output-input combination solves

$$\max_{h^a, h^{\sigma(a)}, L, x} y - w^a h^a - w^{\sigma(a)} h^{\sigma(a)} - zL - x \tag{1}$$

$$\text{s.t.}$$

$$y = F\left(h^a, h^{\sigma(a)}, L, x\right).$$

At this point, we note that profit maximization is arguably a strong assumption, particularly for agricultural household production in developing countries (see, e.g., Udry, 1996). Therefore, in our following analysis we will allow for possible deviations from exact profit maximizing behavior. These deviations may be interpreted as reflecting cross-household variation in production technologies or productive efficiencies.

Second, at the consumption level, we adopt the collective approach of Chiappori (1988, 1992, 1997) and assume that within-household allocations are Pareto efficient. Formally, this means that every matched couple $(a, \sigma(a))$ chooses a consumption allocation that solves

$$\max_{l^a, l^{\sigma(a)}, q_a, q_{\sigma(a)}, Q} U^a\left(l^a, q_a, Q\right) + \mu U^{\sigma(a)}\left(l^{\sigma(a)}, q_{\sigma(a)}, Q\right) \tag{2}$$

$$\text{s.t.}$$

$$w^a l^a + w^{\sigma(a)} l^{\sigma(a)} + q_a + q_{\sigma(a)} + Q \leq w^a T^a + w^{\sigma(a)} T^{\sigma(a)} + n^{(a,\sigma(a))} + \pi^{(a,\sigma(a))},$$

where μ represents the Pareto weight of male $\sigma(a)$ relative to female a, and where $\pi^{(a,\sigma(a))}$ is the profit that results from the profit maximization program (1). Note that Pareto weights are in general not constant. For instance, they will typically vary with wages or marriage market characteristics (such as sex ratios). Attractively, these Pareto weights capture the intrahousehold sharing of resources: a higher value for μ implies that the household decisions reflect to a greater degree male $\sigma(a)$'s preferences.

Third, we assume that the marriage market is stable. Using the definition of Gale and Shapley (1962), marriage stability imposes that marriage matches satisfy the conditions of Individual Rationality and No Blocking Pairs. To formalize the notion of Individual Rationality, let U_H^a and U_H^b represent female a's and male b's utility in their marriage. These utilities follow from the above optimization program. Let us further denote the female's and male's maximum attainable utilities as singles by U_S^a and U_S^b respectively. Note that singles are also consumer-producer households. Their production technologies, however, depend only on their own time spent on agricultural labor, land L and the other input x. Individual Rationality requires

$$U_H^a \geq U_S^a \text{ and } U_H^b \geq U_S^b. \tag{3}$$

Intuitively, Individual Rationality imposes that no female or male wants to exit their marriage and become a single.

Next, to formalize the condition of No Blocking Pairs, we let $U^a_{P_{(a,b)}}$ and $U^b_{P_{(a,b)}}$ represent any possible realization of utilities for female a and male b if they formed a pair. Then, the No Blocking Pair requirement imposes that

$$U^i_{P_{(a,b)}} > U^i_H \text{ implies } U^{i'}_H > U^{i'}_{P_{(a,b)}} \text{ for } i, i' \in \{a, b\}, i \neq i'. \tag{4}$$

In words, a marriage market allocation has No Blocking Pairs if no female a and male b are both better off, with at least one of the two strictly better off, by remarrying each other instead of staying with their current partners.

In what follows, we will quantify deviations from the Individual Rationality and No Blocking Pair conditions by Individual Rationality (IR) and Blocking Pair (BP) indices, which measure the degree of marriage instability. We will compute these indices under the maintained assumptions that intrahousehold consumption allocations are Pareto efficient and production allocations are profit maximizing. As indicated above, we will also show how we can allow for deviations from exact profit maximizing behavior (due to technological heterogeneity or productive inefficiency) in our empirical analysis.

3.3 Marriage stability: empirical conditions

To define our empirical conditions for a stable marriage allocation, we assume a data set $\mathcal{D}$ that contains the following information for a given marriage market:

- matching function σ,
- time uses l^i, m^i and h^i (and time endowment T^i) of each individual i,
- wage w^i of each individual i,
- consumption quantities $(q^{(a,\sigma(a))}, Q^{(a,\sigma(a))})$ of every matched couple $(a, \sigma(a))$,
- land quantities L^a, $L^{\sigma(a)}$ and $L^{(a,\sigma(a))}$ of every matched couple $(a, \sigma(a))$,
- land price $z^{(a,\sigma(a))}$ of every matched couple $(a, \sigma(a))$,
- input quantity $x^{(a,\sigma(a))}$ of every matched couple $(a, \sigma(a))$,
- output quantity $y^{(a,\sigma(a))}$ of every matched couple $(a, \sigma(a))$,
- nonlabor income $n^{(a,\sigma(a))}$ of every matched couple $(a, \sigma(a))$.

We remark that the set $\mathcal{D}$ does not include information on individuals' private consumption; only the aggregate household quantities $q^{(a,\sigma(a))}$ are observed, which is usually the case for household data. The individuals' private quantities will be treated as unknowns in our empirical conditions for marriage stability.[5] Next, in what follows we will assume that wages and land prices remain the same when individuals exit marriage (and become single or remarry), so that divorce has no productivity effects. The assumption that prices and wages are perfectly observed is relaxed below (see Section 3.4).

Characterizing stable marriage. As explained in Section 3.2, we say that the data set $\mathcal{D}$ is consistent with a stable matching if it allows the specification of individual utility functions U^a and U^b that represent the observed consumption behavior as Pareto efficient and the observed marriages as stable. We use revealed preference conditions that are intrinsically nonparametric, in the sense that they do not require an explicit (parametric) specification of the functions U^a and U^b. In particular, we obtain the following testable implications for a stable marriage matching.[6]

Proposition 1 *A* necessary condition *for the data set $\mathcal{D}$ to be consistent with a* ***stable matching*** *σ is that there exist for each matched pair $(a, \sigma(a))$, $a \in A$,*

a. individual quantities $q_a^{(a,\sigma(a))}$, $q_{\sigma(a)}^{(a,\sigma(a))} \in \mathbb{R}_+$ for which $q_a^{(a,\sigma(a))} + q_{\sigma(a)}^{(a,\sigma(a))} = q^{(a,\sigma(a))}$,

b. and nonlabor incomes N^a, $N^{\sigma(a)} \in \mathbb{R}_+$ for which $N^a + N^{\sigma(a)} = n^{(a,\sigma(a))} + x^{(a,\sigma(a))} + z^{(a,\sigma(a))}L^{(a,\sigma(a))}$,

such that the following constraints are met for all females $a \in A$ and males $b \in B$:

i. the individual rationality restrictions

$$\begin{aligned} N^a + z^{(a,\sigma(a))}L^a + w^aT^a &\leq w^al^a + q_a^{(a,\sigma(a))} + Q^{(a,\sigma(a))}, \\ N^b + z^{(\sigma(b),b)}L^b + w^bT^b &\leq w^bl^b + q_b^{(\sigma(b),b)} + Q^{(\sigma(b),b)}, \end{aligned} \tag{5}$$

ii. and the no blocking pair restrictions

$$\begin{aligned} &\left(N^a + z^{(a,\sigma(a))}L^a + w^aT^a\right) + \left(N^b + z^{(\sigma(b),b)}L^b + w^bT^b\right) \\ &\leq \left(w^al^a + w^bl^b\right) + \left(q_a^{(a,\sigma(a))} + q_b^{(\sigma(b),b)}\right) + \max\{Q^{(a,\sigma(a))}, Q^{(\sigma(b),b)}\}. \end{aligned} \tag{6}$$

[5]In our empirical application, part of the private consumption will be assignable to men and women (i.e. individual expenditures on health, education and clothing; see Appendix B). Such information is easy to include in the linear conditions in Proposition 1. It implies lower bound restrictions on the unknowns $q_a^{(a,\sigma(a))}$ and $q_{\sigma(a)}^{(a,\sigma(a))}$. For ease of notation, we do not explicitly consider this refinement here.

[6]See Appendix A for the proofs of our results.

Moreover, a sufficient condition *for the data set $\mathcal{D}$ to be consistent with a stable matching σ is that, in addition, the inequalities* (6) *are strict for $b \neq \sigma(a)$.*

Restrictions (a) and (b) of Proposition 1 specify feasibility constraints that apply to the unknown individual quantities and nonlabor incomes for the matched pairs. These restrictions are associated with the assumption that households choose Pareto efficient intrahousehold allocations. Restrictions (i) and (ii) can be given a "revealed preference" interpretation in terms of a stable marriage allocation. For instance, the inequalities (5) in requirement (i) require, for each individual male and female, that the budget constraints under single status (with income $N^a + z^{(a,\sigma(a))}L^a + w^a T^a$ for female a and $N^b + z^{(\sigma(b),b)}L^b + w^b T^b$ for male b) do not allow buying a bundle that is strictly more expensive than the one consumed under the current marriage (i.e. $\left(l^a, q_a^{(a,\sigma(a))}, Q^{(a,\sigma(a))}\right)$ for female a and $\left(l^b, q_b^{(\sigma(b),b)}, Q^{(\sigma(b),b)}\right)$ for male b). Indeed, if this requirement were not met, then at least one man or woman would be better off (i.e. could attain a strictly better bundle) as a single, which would mean that the marriage allocation is not stable. A similar intuition applies to the inequalities (6) in requirement (ii), which pertain to potentially blocking pairs consisting of females a and males b (see Appendix A for additional explanation).[7]

Some remarks are in order. First, Proposition 1 implies that our structural model is identified. That is, for each solution of the unknown variables that satisfies the empirical constraints in the proposition, we can construct utility functions U^a, U^b and a Pareto weight μ that represents the data in terms of stable marriage allocation. In general, however, the solution to the constraints in Proposition 1 will not be unique, which means that this revealed preference approach typically obtains set identification of the structural components U^a, U^b and μ. We refer to Cherchye, De Rock and Vermeulen (2011) for a detailed discussion on set identification in the context of the collective model of household consumption. These authors also explain the main differences between set identification on the basis of revealed preference characterizations and point identification that is typically pursued in the so-called differential approach to characterizing collective consumption behavior (see, e.g., Chiappori and Ekeland, 2009).

Three further remarks are of a practical nature and pertain to bringing the characterization in Proposition 1 to observational data. First, consistency of $\mathcal{D}$ with a stable matching requires that it is possible to specify individual quantities $q_a^{(a,\sigma(a))}$, $q_{\sigma(a)}^{(a,\sigma(a))}$ and nonlabor incomes $N^a, N^{\sigma(a)}$ that satisfy a set of constraints that are linear in these unknowns. Therefore, a convenient feature of the conditions in Proposition 1 is that they can be checked

[7] We assume that children are captured by the public good, so that these are sufficient conditions for both spouses to be able to afford child custody on divorce. Allowing child custody (and its associated cost) to be spouse-specific would increase the attractiveness of divorce for the spouse who does not receive child custody.

through linear programming, so that they are straightforward to apply in practice. Next, as argued in Cherchye, Demuynck, De Rock and Vermeulen (2017b), the empirical requirement defining the sufficient condition for data consistency with a stable marriage allocation is a very mild one that is easy to verify in practice. Therefore, we will not explicitly discuss this empirical requirement in what follows. Finally, we note that the individual land quantities L^a and L^b are observed in our Malawi data set. This defines natural lower bounds on the left hand sides of the inequality restrictions (5) and (6).

Quantifying marriage instability. An important focus of our empirical analysis is on marriage instability. As explained before, we quantify marital instability in terms of individuals' consumption gains from divorcing and remaining single or remarrying. More specifically, we use our model to define two structural measures of instability: the Individual Rationality (IR) indices capture how much better off (in consumption terms) individuals would be as a single person, and the Blocking Pair (BP) indices measure how much better off individuals would be when remarrying other partners in the same marriage market.

To operationalize these ideas, for each exit option from marriage (i.e. becoming single or remarrying another potential partner), we quantify the minimal within-marriage consumption increase that is needed to represent the observed marriage as stable with respect to the given exit option (as characterized by the conditions (i) and (ii) in Proposition 1). This indicates how far the observed behavior (with the original income levels) is from stable behavior. Conversely, it measures the possible economic gain from divorce when choosing a particular exit option and, therefore, we can interpret it as revealing the degree of marriage instability.

Formally, starting from our characterization in Proposition 1, we include an instability index in each restriction of individual rationality ($s^{IR}_{a,\emptyset}$ for the female a and $s^{IR}_{\emptyset,b}$ for the male b) and no blocking pair ($s^{BP}_{a,b}$ for the pair (a,b)). We replace the inequalities (5) by

$$\begin{aligned}\left(N^a + z^{(a,\sigma(a))}L^a + w^aT^a\right) - s^{IR}_{a,\emptyset} &\leq w^al^a + q_a^{(a,\sigma(a))} + Q^{(a,\sigma(a))},\\ \left(N^b + z^{(\sigma(b),b)}L^b + w^bT^b\right) - s^{IR}_{\emptyset,b} &\leq w^bl^b + q_b^{(\sigma(b),b)} + Q^{(\sigma(b),b)},\end{aligned} \tag{7}$$

and the inequalities (6) by

$$\begin{aligned}&\left(N^a + z^{(a,\sigma(a))}L^a + w^aT^a\right) + \left(N^b + z^{(\sigma(b),b)}L^b + w^bT^b\right) - s^{BP}_{a,b}\\ &\leq \left(w^al^a + w^bl^b\right) + \left(q_a^{(a,\sigma(a))} + q_b^{(\sigma(b),b)}\right) + \max\{Q^{(a,\sigma(a))}, Q^{(\sigma(b),b)}\},\end{aligned} \tag{8}$$

and we add the restriction $0 \leq s^{IR}_{a,\emptyset}, s^{IR}_{\emptyset,b}, s^{BP}_{a,b}$. The indices $s^{IR}_{a,\emptyset}$, $s^{IR}_{\emptyset,b}$ and $s^{BP}_{a,b}$ represent individuals' consumption gains when choosing particular exit options from marriage: $s^{IR}_{a,\emptyset}$ when female a becomes single, $s^{IR}_{\emptyset,b}$ when male b becomes single, and $s^{BP}_{a,b}$ when a and b remarry with each other. Clearly, imposing $s^{IR}_{a,\emptyset}$, $s^{IR}_{\emptyset,b}$, $s^{BP}_{a,b} = 0$ obtains the original (sharp) conditions in Proposition 1, while higher values for $s^{IR}_{a,\emptyset}$, $s^{IR}_{\emptyset,b}$ and $s^{BP}_{a,b}$ correspond to larger deviations from stable marriage behavior. We measure the degree of instability by computing

$$\min_{s^{IR}_{a,\emptyset}, s^{IR}_{\emptyset,b}, s^{NBP}_{a,b}} \sum_a s^{IR}_{a,\emptyset} + \sum_b s^{IR}_{\emptyset,b} + \sum_a \sum_b s^{BP}_{a,b}, \tag{9}$$

subject to the feasibility constraints (a) and (b) in Proposition 1 and the linear constraints (7) and (8). By solving (9), we compute IR indices for the Individual Rationality constraints ($s^{IR}_{a,\emptyset}$ and $s^{IR}_{\emptyset,b}$ in (7)) and BP indices for the No Blocking Pairs constraints ($s^{BP}_{m,w}$ in (8)). Correspondingly, for each exit option, we can define an associated gain from divorce. In our application, we will define "relative" divorce gains by setting out these gains as proportions of current household income. In the next section, we will define a modified version of the objective (9) to address empirical concerns related to unobserved input prices and cross-household heterogeneity in technologies and productive (in)efficiencies.

3.4 Unobserved input prices and production inefficiency

Thus far, we have assumed that prices of the inputs of the household production are observed. In a setting where most households are farmers and only few work off-farm, observed wages are missing or upward biased, while agricultural productivity is more important for economic attractiveness but is not measured in the data. When prices and wages are not observed, shadow prices can be used instead. To obtain shadow prices, we use the structural model that we defined in Section 3.2. In particular, as in Chiappori (1997), we assume profit maximizing behavior under constant returns to scale. In the spirit of Proposition 1, we present a revealed preference characterization, which here means that it does not require an explicit specification of the production technology (represented by the function F).[8] In what follows, we will also show how we can account for deviations from exact profit maximization (because of technological heterogeneity or profit inefficiencies) in our empirical analysis.

Let the true wages (w^i for each individual $i = a, b$) and land prices ($z^{(a,\sigma(a))}$ for each matched pair $(a, \sigma(a))$) be unobserved. Then, we can define shadow wages and prices under the identifying assumption of profit maximizing behavior. Specifically, we say that the data set $\mathcal{D}$ is consistent with shadow profit maximization if we can specify a production function F that represents the observed production behavior as profit maximizing under these shadow

[8] See, for example, Afriat (1972) and Varian (1984) for seminal contributions on this nonparametric approach to analyzing efficient production behavior.

wages and land prices. The following result is an adaptation of Theorem 6 in Varian (1984) to our particular setting.

Proposition 2 *The data set $\mathcal{D}$ is consistent with* ***shadow profit maximization*** *if and only if, for each matched pair $(a, \sigma(a))$ $(a \in A)$, there exist shadow wages w^a, $w^{\sigma(a)} \in \mathbb{R}_+$ and a land price $z^{(a,\sigma(a))} \in \mathbb{R}_+$ that satisfy*

$$0 = y^{(a,\sigma(a))} - \left[w^a h^a + w^{\sigma(a)} h^{\sigma(a)} + z^{(a,\sigma(a))} \left(L^a + L^{\sigma(a)} + L^{(a,\sigma(a))}\right) + x^{(a,\sigma(a))}\right] \tag{10}$$

such that, for all $a' \in A$,

$$0 \geq y^{(a',\sigma(a'))} - \left[w^a h^{a'} + w^{\sigma(a)} h^{\sigma(a')} + z^{(a,\sigma(a))} \left(L^{a'} + L^{\sigma(a')} + L^{(a',\sigma(a'))}\right) + x^{(a',\sigma(a'))}\right]. \tag{11}$$

The restrictions (10) and (11) require that there exist shadow prices such that the observed input-output combination of each matched pair $(a, \sigma(a))$ achieves a profit of zero (see (10)), which must exceed the profit for any household $(a', \sigma(a'))$ (with $a' \in A$) under the same prices (see (11)). Note that this condition of zero maximum profit directly follows from our constant returns to scale assumption. We can append these profit efficiency restrictions to the stability conditions above. As a result, our marriage stability analysis will use shadow wages and land prices that are identified under the assumption of profit maximizing household production. See also the linear program that we present below in (14).

Our empirical analysis will make use of two extensions of the characterization in Proposition 2. First, the characterization only imposes that shadow prices should be non-negative. Obviously, this allows for shadow prices that are unrealistic proxies of the true (unobserved) prices (e.g., prices that are infinitely high). To exclude such unrealistic scenarios, we impose lower and upper bounds on possible prices. Specifically, we append the restrictions

$$\underline{w}^a \leq w^a \leq \overline{w}^a,\ \underline{w}^b \leq w^b \leq \overline{w}^b \text{ and } \underline{z}^{(a,\sigma(a))} \leq z^{(a,\sigma(a))} \leq \overline{z}^{(a,\sigma(a))},$$

where $\underline{w}^a$, $\underline{w}^b$, $\underline{z}^{(a,\sigma(a))} \in \mathbb{R}_{++}$ and $\overline{w}^a$, $\overline{w}^b$, $\overline{z}^{(a,\sigma(a))} \in \mathbb{R}_{++}$ are predefined lower and upper bounds. Appendix B explains how we derive these bounds from the observed price information in our data set.

Our second extension pertains to the fact that the characterization in Proposition 2 implicitly assumes that different households are exactly profit efficient and characterized by a homogeneous production technology (defined at the marriage market level). Clearly, in practice we need to account for unobserved heterogeneity in technologies and productive (in)efficiencies across households (see, e.g., Udry, 1996). To do this, we introduce deviational variables π^{a+}, π^{a-}, $\pi^{a,a'} \in \mathbb{R}_+$ for each matched pair $(a, \sigma(a))$. These variables capture possible deviations from the original (sharp) conditions in Proposition 2, which can thus be explained as deviations from exact profit maximization under a homogeneous production technology.[9]

Formally, in our profit characterization in Proposition 2, we replace the equality restriction (10) by

$$\begin{aligned} \pi^{a+} - \pi^{a-} = y^{(a,\sigma(a))} - \\ \left[w^a h^a + w^{\sigma(a)} h^{\sigma(a)} + z^{(a,\sigma(a))}\left(L^a + L^{\sigma(a)} + L^{(a,\sigma(a))}\right) + x^{(a,\sigma(a))}\right], \end{aligned} \tag{12}$$

and the inequality restriction (11) by

$$\begin{aligned} \pi^{a,a'} \geq y^{(a',\sigma(a'))} - \\ \left[w^a h^{a'} + w^{\sigma(a)} h^{\sigma(a')} + z^{(a,\sigma(a))}\left(L^{a'} + L^{\sigma(a')} + L^{(a',\sigma(a'))}\right) + x^{(a',\sigma(a'))}\right]. \end{aligned} \tag{13}$$

The variables π^{a+}, π^{a-}, $\pi^{a,a'}$ account for deviations from the zero maximum profit that appears on the left hand side in the original conditions (10) and (11). That is, they capture deviations from the assumption of profit maximizing behavior under constant returns to scale with a homogeneous household technology.

In our application, we use shadow prices that minimize the aggregate value of the deviational variables, $\sum_a \left(\pi^{a+} + \pi^{a-} + \sum_{a'} \pi^{a,a'}\right)$. This implies that we replace the objective (9) defined above by (with $0 \leq \alpha \leq 1$)

$$\begin{aligned} \min_{s^{IR}_{a,\emptyset}, s^{IR}_{\emptyset,b}, s^{NBP}_{a,b}, \pi^{a+}, \pi^{a-}, \pi^{a,a'}} \alpha \left(\sum_a s^{IR}_{a,\emptyset} + \sum_b s^{IR}_{\emptyset,b} + \sum_a \sum_b s^{BP}_{a,b}\right) \\ + (1-\alpha)\left(\sum_a \left(\pi^{a+} + \pi^{a-} + \sum_{a'} \pi^{a,a'}\right)\right), \end{aligned} \tag{14}$$

subject to the constraints (a) and (b) in Proposition 1, the stability constraints (7) and (8) and the profit maximization constraints (12) and (13). Because all constraints are linear in unknowns, we can compute the solution values of $s^{IR}_{a,\emptyset}$, $s^{IR}_{\emptyset,b}$, $s^{BP}_{a,b}$, π^{a+}, π^{a-} and $\pi^{a,a'}$ by straightforward linear programming. Summarizing, the above minimization program looks for optimal feasible values of unobserved individual quantities, shadow prices (including

[9] Deviational variables are also used in the "goal programming" approach to deal with infeasible linear programs.

wages) and nonlabor incomes in such a way that deviations from stability and profit maximization are minimized.

In (14), the parameter α is a tuning parameter that represents the "penalization" weight of the marriage instability indices relative to the technological heterogeneity variables. Since we use profit maximizing behavior as our identifying assumption for shadow wages and land prices, we set α to be very small.[10] This can be thought of as a two-stage optimization process: in the first stage, we define shadow prices as the prices that correspond to minimal deviations from our profit maximization conditions (measured by $\sum_a \left(\pi^{a+} + \pi^{a-} + \sum_{a'} \pi^{a,a'}\right)$); in the second stage, we compute instability indices for the given shadow prices (by minimizing $\sum_a s^{IR}_{a,\emptyset} + \sum_b s^{IR}_{\emptyset,b} + \sum_a \sum_b s^{BP}_{a,b}$).

4 Data

Our data are drawn from the third Malawi Integrated Household Survey (IHS). We use the baseline survey conducted in 2010 and the second wave in 2013, where approximately one quarter of households were re-interviewed. These households were chosen randomly, and both the baseline sample and the panel subsample were designed to be nationally representative of the population of Malawi.[11] We restrict our sample to rural, monogamous households that engage in agriculture.[12] This yields a sample of 8624 households in 2010, of which 5924 were married. Of the married households, approximately one third ($N = 1404$) are observed three years later. We allow singles to form potential blocking pairs with married individuals, but our instability indices are only estimated for married individuals. Appendix B discusses the construction of the dataset in more detail.

A crucial component of our analysis is the specification of marriage markets, within which individuals can form potential blocking pairs. As stated earlier, marriages tend to be local in Malawi. In the IHS dataset, approximately 45% of married individuals are from the village they live in, while a further 25% are from another village within the same district. We use this fact to guide our definition of marriage markets. In particular, we use the GPS coordinates of villages to construct clusters of two to three geographically close villages, which form a marriage market. We use the k-means unsupervised machine learning algorithm, which partitions the data into k clusters using the squared Euclidean distance. We set the number

[10] Specifically, we use $\alpha = 10^{-6}$ in our following empirical application. We also experimented with alternative values for α (with α far below $(1-\alpha)$) but this did not affect our main conclusions.

[11] In the baseline survey, 768 communities were selected based on probability proportional to size, within which 16 households were randomly sampled.

[12] We use survey weights in all our descriptive statistics and also take into account the fact that the primary sampling units are villages by clustering at the village level.

of clusters to 300, so that the number of households per cluster ranges from 5 to 58, with the average number of households per cluster at 33.5. The fact that we construct small marriage markets based on geographically proximate villages increases the likelihood of encounters between individuals in these marriage markets. As households are randomly sampled at the village level, the sample will be representative of the *types* of individuals in a person's marriage market. In this sense, although we do not observe the complete population of each marriage market, we observe a representative subset of types. We are implicitly assuming that the remarriage market is captured by these geographical clusters; thus, it cannot be the case that individuals only remarry people in faraway villages, for example, due to social stigma. Indeed, the social stigma of divorce is likely to be fairly low in this setting, given the high divorce rate. Finally, the more individuals there are in the marriage market, the more likely that there is a profitable new match. Thus, the size of the marriage market can affect the values of marital instability and we address this by controlling for marriage market fixed effects in our empirical analysis of divorce decisions (see Section 6).

Table 2 describes the characteristics of our sample. On average, the household head is middle-aged and 76% of household heads have no education. The average household has approximately three children and almost two acres of land. Most consumption is non-assignable, with 23% of consumption devoted to public goods and 2% devoted to the man's and woman's assignable goods, on average. The primary component of non-assignable consumption is food, which forms 64% of total consumption, on average. Clothing forms 3% of annual consumption, while public consumption includes utilities and house-related expenses, which form 14% of annual consumption, on average. All spending on children (education, health, clothing) is subsumed in public consumption. Thus, the majority of our households' budget is spent on food, with a further large share spent on housing and utilities.

Table 2: Summary statistics of rural, monogamously married households in 2010 IHS

Variable	Mean	Standard error
Age of head	40.39	(0.22)
Head has no education (0-1)	0.76	
Head has primary education (0-1)	0.10	
Head has secondary education (0-1)	0.12	
Head has tertiary education (0-1)	0.01	
Number of children	2.95	(0.03)
Land (acres)	1.94	(0.04)
Total consumption ('000s)	210.70	(3.55)
Public share of consumption	0.23	(0.00)
Private share of consumption, woman	0.01	(0.00)
Private share of consumption, man	0.01	(0.00)
Nonassignable share of consumption	0.75	(0.00)
Number of observations (N)	5924	
Number of marriage markets	300	

5 Estimation results of the structural model

In this section, we discuss the estimation results from the structural model. We estimate our model using the first wave of the survey (2010), and reserve the second wave of the panel (2013) for our out-of-sample prediction of divorces. In particular, the optimization program in equation (14) in Section 3.3 yields several outputs: instability indices for each possible pair in each marriage market; instability indices for each individual for the outside option of being single; shadow wages; shadow land prices; individual nonlabor income; intrahousehold sharing of consumption; and deviational variables, which represent the proximity to profit maximizing behavior. We present the estimated wages, land prices and nonlabor income in Table 3, while Table 4 displays the instability indices. As profit maximization is not the focus of our analysis, we do not report the deviational variables; in any case, the averages are very close to zero, which reflects that we set the tuning parameter α very close to zero in equation (14).

We find that, on average, women have a significantly lower shadow wage than men, which is consistent with reported non-agricultural wages in the survey. Women also have significantly lower land income than men, on average, which is partly driven by the fact that the average woman owns less land than the average man. Nonlabor income is overall high for both men and women, and is defined as the shortfall between income and consumption and leisure, so that high nonlabor income is driven by high leisure, low agricultural productivity, low land price and small landholdings. In particular, reported leisure is very high in the survey, suggestive of overreporting, and is the most important contributor to the large average nonlabor income. These observations are useful to bear in mind when we discuss our estimates of the relationship between the outputs of our structural model and divorce in Section 6.2.1.

Table 3: Summary statistics of wages, land income and nonlabor income estimated in structural model

Variable	Men		Women	
	Mean	Standard error	Mean	Standard error
Wage	124.09	(0.878)	117.32	(0.784)
Land income ('000s)	9.001	(0.356)	4.110	(0.166)
Nonlabor income ('000s)	137.82	(3.187)	206.30	(2.771)
Number of observations			5924	

Next, we describe the estimated value of divorce from our model in Table 4. For each individual, we define two Blocking Pair (*BP*) indices: the *BPmax* index represents the individual's gain associated with the most attractive remarriage option, and the *BPavg* index gives the individual's average gain from remarriage, across all possible new pairs that this individual could form in their marriage market. The Individual Rationality (*IR*) index measures the gain from divorcing and being single. All indices are expressed relative to the household's total income.

Table 4: Summary statistics of instability indices estimated in structural model

	Men			Women		
	Mean	Standard error	% Non-zero	Mean	Standard error	% Non-zero
BPmax	0.716	(0.072)	16.85	3.432	(0.171)	64.89
BPavg	0.253	(0.022)	16.85	0.128	(0.007)	64.89
IR	1.914	(0.118)	47.42	0	*N/A*	0
Number of observations				5924		

Some interesting observations emerge. First, we estimate that 65% of women have a profitable match in their marriage market, while fewer than 17% of men have a profitable match. In contrast, no women in our sample would prefer to be single over staying married, while over 47% of men would prefer the single option. From the *BPmax* estimates we learn that, on average, women gain more by choosing the most attractive remarriage option than men. However, our *BPavg* results reveal that women's gains from selecting the "average" remarriage possibility are generally lower than men's. This implies that women have many unattractive potential matches and some very attractive potential matches, while men have mostly mediocre, somewhat attractive matches. We sharpen the intuition for our instability concepts in Appendix C, where we provide an example of the instability network of one marriage market in our sample.

The model predicts that almost half of the men in our sample would like to be single, while more than half of the women have profitable remarriage opportunities. In the context of a frictionless marriage market, this implies that the model omits unobserved costs of being single for men and remarriage for women. Plausibly, for men, there may be an unobserved benefit to being married, such as the domestic labor of their wives, while for women, there may be an unobserved cost of divorcing and remarrying, such as social stigma.

At this point, we note that the absence of domestic non-agricultural labor, which is currently subsumed in leisure both in the model and the data, can explain the finding that no woman would prefer to be single, as virtually all domestic labor in Malawi is carried out by women. This means that women who engage in many hours of domestic work appear to have more leisure than they actually do. As a result, their outside option of being single appears less attractive. If data on domestic labor were available, this would reduce women's leisure and make it more likely that some of them would prefer to be single.

Turning next to summary statistics of divorce between 2010 and 2013, Table 5 shows that 11.7% of households divorce between the two waves of the survey.[13] Of those women with known marital status in 2013, there is a similar number of single women and remarried women, while most men remarry. Finally, Table 6 compares the characteristics of couples who divorce with those who do not. We find that both men and women who divorce have higher values of all instability indices in 2010, and we present a rigorous analysis of this relationship in Section 6.2. The table also shows that households who divorce have significantly lower total consumption, fewer children and less land. Among couples who are still married, the household head is older, on average, in line with standard intuition that poor matches are dissolved early on.

[13] There are some divorced households in 2013 where one of the spouses could not be re-interviewed; this is why the total number of divorced men or women with known marital status is fewer than the total number of divorced households.

Table 5: Changes in marital status between 2010 and 2013, Malawi IHS

N (%)	Married	Divorced - remarried	Divorced - single	Total
Couples	1240	164 (11.7%)		1404
Women	1240	74 (5.4%)	64 (4.6%)	1378
Men	1240	84 (6.2%)	21 (1.6%)	1345

Table 6: Summary statistics of characteristics of couples who divorce and do not divorce between 2010-2013

	Divorce		Do not divorce	
	Mean	Standard error	Mean	Standard error
BPmax, woman	3.81	(0.49)	3.31	(0.32)
BPmax, man	0.72	(0.29)	0.59	(0.13)
BPavg, woman	0.14	(0.02)	0.12	(0.01)
BPavg, man	0.33	(0.20)	0.22	(0.04)
IR, man	1.94	(0.40)	1.74	(0.25)
Age of head	35.04	(1.44)	40.83	(0.55)
Number of children	2.49	(0.16)	3.13	(0.06)
Land (acres)	1.72	(0.16)	2.06	(0.07)
Total consumption ('000s)	203.29	(12.51)	237.04	(9.44)
Number of observations	164		1240	
Number of marriage markets	117			

6 Divorce and the marriage market

In this section, we demonstrate the empirical relevance of our structural model by showing that our structurally defined instability indices are correlated with individual and household characteristics that are plausible measures of individuals' outside options in the data, and that they predict future divorce and remarriage.

6.1 What drives instability?

We show that our instability indices are correlated with measures of own match quality and the characteristics of the marriage market. Recall that household characteristics were not used in the estimation of our model through the linear program in equation (14), which only relied on information about aspects of consumption and production. Therefore, it is a valuable exercise to explore whether the instability indices correlate with household characteristics in the way we would expect. In particular, we estimate the correlation between the instability indices and the age and education of the spouses, the number of children they have, as well as dummy variables for whether they have the same age and the same education (intended to capture the value of assortative mating). We also include characteristics of the marriage market, including the number of churches (to capture religiousness), the distance to the nearest road, the distance to the nearest urban centre, as well as the number of households in the marriage market. The equation we estimate is

$$s^j_{i,m} = \alpha_0 + \boldsymbol{\alpha}_1 \mathbf{H}_{i,m} + \boldsymbol{\alpha}_2 \mathbf{X}_m + \varepsilon_{i,m},$$

where $s^j_{i,m}$ is the instability index j ($j = BPmax, BPavg, IR$) of individual i living in marriage market m, $\mathbf{H}_{i,m}$ are characteristics of individual i's household, and $\mathbf{X}_m$ are characteristics of individual i's marriage market; see the estimates in Table 11 in Appendix D.

We find that the more educated the household head (which is the husband in virtually all cases), the lower are the wife's *BP* indices (i.e. her remarriage possibilities are less attractive). This estimated effect is monotonically increasing in the education level of the household head. For example, a woman living in a household where the head has primary school education has an average *BPmax* index that is 57 percentage points lower than a comparable woman where the head has no education. Recall that the *BP* indices are defined relative to household income, so that this coefficient captures a decline in the ratio of 0.57. On the other hand, the education of the household head is not correlated with the *BP* indices of the husband. Instead, we find that he has stronger outside options when he is older (likely because this is correlated with accumulated wealth), but weaker outside options when he has a younger wife (as this is correlated with a wife's fecundity and hence attractiveness). Children in marriage significantly reduce the value of all outside options. Next, we observe a relationship between connectedness and stability: marriage markets that are far away from roads and urban centres are more stable. A one kilometer increase in the marriage market's distance to the nearest road reduces the wife's average *BPmax* by 4.2 percentage points, while the same increase to the nearest urban centre reduces this index by 0.2 percentage points on average. Finally, as expected, larger marriage markets are associated with larger average values of the *BP* indices while having no estimated effect on the *IR* index.

6.2 Divorce

6.2.1 Main results

We now present the empirical analysis of divorce. In particular, we estimate whether our structural measures of the value of the remarriage market, and of being single, can predict future divorces. Note that there is no *a priori* reason to expect an empirical association between our measures of instability and future divorces, as no information from the 2013 wave was used in the estimation of the structural model: hence, this provides an out-of-sample test of our model. We estimate a linear model of divorce between 2010 and 2013, with the BP indices of the spouses and the IR index of the husband as covariates (recall that the IR index is zero for all wives). We include marriage market fixed effects, and also control for all household-level variables reported in Table 11, as they covary with the instability indices and potentially also with divorce probability. The equation we estimate is

$$d_{h,m} = \beta_0 + \beta_1 s^j_{i,h,m} + \beta_2 s^j_{i',h,m} + \boldsymbol{\beta}_3 \mathbf{H}_{h,m} + \boldsymbol{\beta}_4 \boldsymbol{\mu}_m + \epsilon_{i,m},$$

where $d_{h,m}$ is a dummy variable that equals one if household h in marriage market m divorces between 2010-2013, and zero if they remain married, $s^j_{i,h,m}$ is the instability index j of spouse i in household h in marriage market m, $s^j_{i',h,m}$ is the instability index j of spouse i', $\mathbf{H}_{h,m}$ are household characteristics and $\boldsymbol{\mu}_m$ are marriage market fixed effects.[14] We estimate these equations separately for $j = BPmax$ and $j = BPavg$, but include $j = IR$ in both of these equations. The estimates are reported in Table 7.

[14] Marriage market fixed effects will capture characteristics that matter for overall divorce propensity, such as the type of descent practiced, the size of the marriage market, and distance to large towns, for example.

Table 7: OLS regressions of divorce between 2010-2013 on instability indices in 2010 and other control variables

	(1)	(2)	(3)	(4)
		Divorced in 2013		
BPmax (woman)	0.014**	0.014**		
	(0.006)	(0.006)		
BPmax (man)	0.001	0.001		
	(0.030)	(0.030)		
IR (man)	0.837	0.848	0.088	0.094
	(2.423)	(2.414)	(1.608)	(1.605)
BPavg (woman)			0.489***	0.491***
			(0.144)	(0.148)
BPavg (man)			0.017	0.017
			(0.025)	(0.025)
N			1406	
R^2	0.126	0.126	0.129	0.130

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head and the number of children the household had in 2010. Columns (2) and (4) also control for dummy variables indicating whether the couple are within two years of age of each other, and whether they have the same level of education.

We find that the instability indices have significant predictive power for future divorce, particularly for measures of the value of the wife's remarriage market. In regression (1), a one percentage point increase in the wife's maximum gain from remarriage raises the probability of divorce by 1.4 percentage points on average. This is a sizeable effect, as the annual divorce probability is approximately 8.5%. In regression (2), we control for measures of assortative mating: dummy variables that equal one if the spouses have the same education level, and the same age (±2 years). The coefficient on *BPmax* is unchanged. In regression (3), we repeat the first specification (without assortative mating variables) but replace *BPmax* with *BPavg*, and we find that a one unit increase in the average remarriage gain for the wife, as a proportion of her household's income, raises divorce probability by 48.9 percentage points.

Note that the impacts of a unit change in the maximum and average gains from remarriage on divorce probability are not directly comparable to each other, as the levels of *BPmax* and *BPavg* are different (see Table 4). Overall, we find that our measures of the value of the woman's remarriage market are predictive of divorce.

We find no significant associations between the measures of husbands' outside options, and divorce. Thus, economic gains to divorce and remarriage matter for the women in our sample, but not the men. In Malawi, women marry young and divorce often. In the demographic literature, it has been argued that women in Malawi use marriage and divorce to improve their economic situation; for example, women are less likely to live in their husband's village in higher order marriages, which is considered empowering (Reniers, 2003). Men, on the other hand, are more likely to remarry younger women, and hence plausibly are motivated by the fecundity, rather than economic circumstance, of their potential spouse. This intuition can explain why we observe that the economic value of a woman's marriage market matters, while that of the man's does not.

We consider and refute alternative explanations that could explain our empirical results. First, note that the absence of domestic (non-agricultural) labor in the data cannot explain the significant estimated effect of the wife's instability index on divorce. As domestic labor is currently subsumed in leisure, marriages appear to be more attractive than they actually are. Consider a woman who engages in a substantial amount of domestic labor: she appears to be in a stable marriage, but at the same time may be unhappy because she works hard, as a result of which she is more likely to divorce. An increase in domestic labor increases stability in our model but at the same time is likely to increase the probability of divorce. Therefore, it cannot explain the positive relationship between the instability indices and divorce probability. Second, we consider that the measure *BPmax* is likely to be sensitive to who is sampled from the marriage market, more so than *BPavg*. An alternative measure that captures the top end of the remarriage distribution, but that is less sensitive to sampling, is the 95th percentile of an individual's *BP* indices (*BP95*). These results are displayed in Table 13 in Appendix E and the coefficients on *BP95* are consistent with those on *BPmax* in Table 7; a one unit increase in *BP95* increases divorce propensity by 4.5 percentage points, which predictably lies between the coefficients on *BPavg* and *BPmax*.

In Table 12 in Appendix E, we estimate the specifications in Table 7 using a logit regression model, where we report marginal effects at means. The marginal effect of *BPavg* is similar in magnitude to the average affect in Table 7 although we lose some precision, while the estimated effect of *BPmax* is virtually unchanged and significant at the 1% level. Next, the inclusion of a dummy for the existence of polygamy in the village does not affect the significant estimated effect of the wife's instability index on divorce, but we do find that the existence of polygamy increases the overall probability of divorce (results not reported for compactness).

As a following exercise, we show that the estimated effect of the *BP* indices cannot be explained by accounting for characteristics such as wages or landholdings. To show this, we control for the *components* of the *BP* indices in the divorce equation. These components are specified in Equation (8) in Section 3.3, and they include wages, land income, nonlabor income, and the intrahousehold sharing of consumption. The structural model estimates an intrahousehold sharing of consumption that is consistent with the *BP* indices, which are set identified and thus not point identified. Hence, we only control for the first three of these components. Clearly, some of these variables are "bad controls", in the sense that they themselves are determined by choices (Angrist and Pischke, 2009). However, we do not interpret this exercise in a causal way. Rather, our goal is to show how the relationship between our instability indices and divorce propensity depends on the *BP* components. In particular, we find that it is not sensitive to the inclusion of these variables at all, from which we conclude that the intrahousehold sharing of consumption is the key (omitted) determinant.

We begin by controlling for the individuals' estimated individual nonlabor incomes and the shadow wages and shadow land income; these results are in specifications (1) and (2) in Table 8. As in the main results, these regressions include marriage market fixed effects and household characteristics. The impact of the wife's *BPavg* and *BPmax* indices on divorce propensity is similar to the main results in Table 7, with both coefficients larger in absolute magnitude. Next, in specifications (3) and (4), we also add this same information, but for the individuals' potential partners: the average nonlabor income, wages and land income of everyone except the household in the marriage market. The impact of the inclusion of these variables on the main coefficients is negligible. In other words, the estimated effect of the instability indices on divorce cannot be explained by a linear combination of own and others' wages, land income and nonlabor income. In this sense, we argue that the estimated effect of the value of remarriage options on divorce is driven by the intrahousehold sharing of consumption.

Table 8: OLS regressions of divorce between 2010-2013 on instability indices in 2010, control variables and wages, nonlabor income and land income from the structural model

	(1)	(2)	(3)	(4)
		Divorced in 2013		
BPmax (woman)	0.019***	0.017**		
	(0.007)	(0.007)		
BPmax (man)	0.004	0.006		
	(0.030)	(0.031)		
IR (man)	1.062	0.656	0.524	0.172
	(2.507)	(2.604)	(1.715)	(1.801)
BPavg (woman)			0.612***	0.578***
			(0.166)	(0.173)
BPavg (man)			0.018	0.020
			(0.026)	(0.026)
N			1406	
R^2	0.132	0.136	0.136	0.140

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head and the number of children the household had in 2010, and the following estimated variables from the structural model: husband and wife's wages, nonlabor income and land income. Columns (2) and (4) also control for the average of these structural variables in the marriage market, excluding the value of the individuals in the household.

6.2.2 Divorce: Remarriage or remaining single?

An important implication of the way that the instability indices are defined is that the *BP* index measures the attractiveness of a potential new match in the marriage market, while the *IR* index measures the attractiveness of being single. Therefore, we should observe these associations in the data. In order to show that this is the case, we estimate the relationship between the *BP* indices and two separate outcomes: divorce and remarriage, and divorce and remaining single. In particular, we define two indicator variables: the variable Remarried takes the value one if an individual divorced and remarried between 2010 and 2013, and zero otherwise (including if they remained married), while Single takes the value one if the individual divorced but was not remarried in 2013, and zero otherwise. We observe this information for most but not all individuals in the survey, and are able to construct these variables separately for men and women. As in the main estimates in Table 7, we control for marriage market fixed effects and household characteristics. The results are in Table 9. For compactness we report the estimated effect of *BPmax* here; the same regressions with *BPavg* are in Table 14 in Appendix E.

Table 9: OLS regressions of marital status in 2013 on instability indices in 2010 and other control variables

	(1)	(2)	(3)	(4)
	Women		Men	
	Remarried	Single	Remarried	Single
BPmax (woman)	0.011***	0.006*	0.009*	0.003
	(0.003)	(0.004)	(0.005)	(0.003)
BPmax (man)	0.004	0.005	0.009	-0.005
	(0.012)	(0.019)	(0.021)	(0.007)
IR (man)	0.012	0.454	0.071	0.650
	(1.073)	(1.471)	(1.629)	(0.454)
N	1380	1380	1347	1347
R^2	0.113	0.105	0.119	0.094

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions. The dependent variable Remarried takes a value of one if the person has divorced and remarried by 2013, and zero otherwise. The variable Single takes a value of one if a person has divorced but has not remarried by 2013, and zero otherwise. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head and the number of children the household had in 2010, and dummy variables for whether the couple are within two years of age of each other, and whether they have the same level of education.

The results are consistent with the premise that the *BP* indices measure the attractiveness of the remarriage market. In particular, a higher value of the wife's *BP* index is associated with a significantly higher probability that the wife divorces and remarries in the next three years, instead of remaining married (regression (1)). The index also predicts divorcing and being single (regression (2)); however, the coefficient is around one half of the size of the coefficient in regression (1). Indeed, not everyone who divorces with the intention of remarrying will manage to do so. Interestingly, the wife's *BP* index is also predictive of the husband divorcing and remarrying, consistent with men preferring to remarry rather than remain single, even if they did not trigger the divorce. The husband's *IR* index does not affect the probability of either divorce status, which is consistent with its weak significance in Table 7, suggesting that the *IR* index does not fully capture the gains to being single. In terms of magnitudes, we find that a one unit increase in the wife's best gain from remarriage, as a proportion of household income, raises the probability that the wife has divorced and remarried, relative to all other marital states, by 1.1 percentage points. It also increases the probability of being divorced and single by 0.6 percentage points, and raises the probability of the husband having remarried by 0.9 percentage points. These magnitudes are similar to those in Table 7.

In Appendix E, we present the estimates of a multinomial logit model of marital status in 2013 (see Tables 15 and 16). Consistent with the OLS results, we find that an increase in the wife's *BP* index is associated with increased odds of divorcing and remarrying by 2013 for both the husband and wife. In particular, a one unit increase in *BPmax* is associated with a 7% higher risk of the woman and 12% higher risk of the man being divorced and remarried, compared to the base category of remaining married. Additionally, an increase in the husband's *BP* index is associated with higher odds of the husband divorcing and remarrying. Neither *BP* index is associated with significantly changed odds of divorcing and being single, compared to remaining married.

6.2.3 Interactions between the remarriage market and other drivers of match quality

As a further exercise, we explore the role of other drivers of marital surplus in divorce, and how they interact with the estimated effects of our economic measures of the remarriage market. We focus on other drivers of match quality and attractiveness that have been well-documented in the literature: age, education, and assortative mating in these factors (see, for instance, Browning, Chiappori and Weiss, 2014, for a recent overview). We have already controlled for these measures in the main results; here, we explore heterogeneity of our main effects with respect to these variables. In this sense, we go some way towards characterizing match quality as consisting of both an economic value, as captured by our *BP* indices, and value from non-economic characteristics. We expect that characteristics that improve the value of the current marriage, such as the number of children, will reduce the predictive impact of *BP* indices on divorce, as these characteristics can compensate spouses for lower "economic attractiveness". We also explore heterogeneity of the main effects with respect to the local sex ratio (defined as the ratio of males over females in a given village, hence exploiting variation between villages within a marriage market). The results for *BPmax* are in Table 10; similar estimates for *BPavg* can be found in Table 17 of Appendix E.

We find precisely estimated differences in the gradient of *BPmax* with respect to age, having the same age as the spouse, and the sex ratio. In particular, the estimated effect of the wife's *BP* index is decreasing with her age, suggesting that being older makes it more difficult to find an alternative partner. Interestingly, we find a significant estimated negative interaction between the husband's *BPmax* and being of the same age, suggesting that husbands value assortative mating on age. In regression (3), we examine the interaction between the sex ratio and the estimated effect of remarriage options on divorce. For a sex ratio equal to one, an increase in the wife's *BPmax* index of one unit increases the probability of divorce by approximately 7.7 percentage points. The more men there are, relative to women, the stronger the estimated effect of the wife's potential gains from remarriage on divorce probability. This is a rational response: if there are more men relative to women in the population, the likelihood of a profitable remarriage is greater.

Finally, we summarize the other, less precisely estimated effects. The interaction term between the number of children and the spouses' *BP* indices is negative, suggesting that having more children reduces the attractiveness of other outside options. This is consistent with the observation that divorce occurs less among couples who have children. Similarly, the coefficient on the interaction between having the same education level, and the *BP* indices, is negative, which suggests that assortative mating on education can compensate for a lack of economic attractiveness.

7 Conclusion

Divorce is a widespread phenomenon with potentially large welfare effects on all parties that are involved. The study of divorce in the economic literature has been largely dominated by the role of economic shocks (with the exception of studies that link intrahousehold choices to divorce decisions, such as Voena, 2015). We argue that the marriage market has a crucial role to play in the decision to divorce. We have defined structural measures of individuals' outside options on the marriage market and shown that they are significant (out-of-sample) predictors of future divorces. These measures are based on a collective model with consumption and agricultural production embedded in a marriage market. We quantify marital instability in terms of Individual Rationality (*IR*) and Blocking Pair (*BP*) indices, which capture spouses' consumption gains to remarrying another individual in the same marriage market (*BP* index) and to being single (*IR* index).

Table 10: OLS regressions of divorce between 2010-2013 on instability indices in 2010 and their interactions with other variables

	(1) *Divorced in 2013*		(2) *Divorced in 2013*		(3) *Divorced in 2013*
BPmax (w)	0.039***	*BPmax* (w)	0.014**	*BPmax* (w)	-0.051
	(0.010)		(0.006)		(0.034)
Age**BPmax* (w)	-0.001***	Same age**BPmax* (w)	0.006	Sex ratio**BPmax* (w)	0.077*
	(0.000)		(0.010)		(0.040)
BPmax (m)	0.014	*BPmax* (m)	0.002	*BPmax* (m)	0.012
	(0.036)		(0.030)		(0.045)
Age**BPmax* (m)	-0.000	Same age**BPmax* (m)	-0.020**	Sex ratio**BPmax* (m)	-0.011
	(0.000)		(0.010)		(0.035)
N	1406		1406		1406

	(4)		(5)
BPmax (w)	0.015	*BPmax* (w)	0.020***
	(0.009)		(0.008)
Same educ**BPmax* (w)	-0.001	# Children**BPmax* (w)	-0.002
	(0.008)		(0.002)
BPmax (m)	-0.011	*BPmax* (m)	0.012
	(0.030)		(0.032)
Same educ**BPmax* (m)	-0.001	# Children**BPmax* (m)	-0.004
	(0.008)		(0.002)
	1406		1406

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head and the number of children the household had in 2010, and dummy variables for whether the couple are within two years of age of each other, and whether they have the same level of education. They also include the variable that is being interacted with the indices (e.g. Sex ratio in regression (3)).

We estimate this model on data drawn from a household survey in Malawi, which has rich information on consumption and production, as well as information on marital status changes over three years. Our key results are as follows. We find that a 1 percentage point increase in the wife's most attractive outside option, relative to her household income, is associated with a 1.4 percentage point higher probability of divorce over the following three years, and increases the probability that she has divorced and remarried by 1.1 percentage points. We find no significant associations between the value of the husband's remarriage market and subsequent divorce, which is consistent with men and women valuing economic characteristics in their partners to different extents (Reniers, 2003). The estimated relationship between the wife's remarriage market and divorce cannot be explained by a linear combination of wages, nonlabor income and land income, indicating that intrahousehold sharing of consumption is the key driver of this relationship. Finally, we find that this estimated effect interacts with other characteristics that affect match quality. In particular, it is dampened by the age of the spouses, and by a shortage of men, relative to women, in the marriage market.

Our findings show that divorce in Malawi is driven, at least partly, by the economic considerations of spouses. In addition, our empirical results validate the set-up of our theoretical model, akin to an out-of-sample test. More generally, they show the value-added of adopting a Beckerian approach that analyses marriage decisions through the lens of a structural model of household decision making. Further, as agricultural productivity is a key determinant of outside options for households reliant on production, our model is applicable to other contexts as well. Finally, the estimation of our model on a longer panel dataset would be informative in observing changes in the value of own marriages and the remarriage market, hence shedding light on which of the two is a more important driver in divorce decisions; we leave this exercise to future work.

References

[1] Afriat, S. "The Construction of Utility Functions from Expenditure Data." *International Economic Review* 8 (1967): pp. 67-77.

[2] Afriat, S. "Efficiency Estimation of Production Functions." *International Economic Review* 13 (1972): pp. 568-598.

[3] Apps, P. and R. Rees. "Labor Supply, Household Production and Intra-Family Welfare Distribution", *Journal of Public Economics*, 60 (1996): pp. 199-219.

[4] Becker, G. S. "A Theory of Marriage: Part I." *Journal of Political Economy* 81 (1973): pp. 813-846.

[5] Becker, G. S. "A Theory of Marriage: Part II." *Journal of Political Economy* 82 (1974): pp. S11-S26.

[6] Becker, G. S., E. M. Landes and R. T. Michael. "An Economic Analysis of Marital Instability." *Journal of Political Economy* 85 (1977): pp. 1141-1187.

[7] Bignami-Van Assche, S., A. Van Assche, P. Anglewicz, P. Fleming and C. van de Ruit. "HIV/AIDS and Time Allocation in Rural Malawi." *Demographic Research* 24 (2011): pp. 671-708.

[8] Blundell, R., P.-A. Chiappori, T. Magnac, and C. Meghir. "Collective labor supply: heterogeneity and nonparticipation." *Review of Economic Studies* 74 (2007): pp. 417-445.

[9] Boheim, R. and J. Ermisch. "Partnership Dissolution in the UK - the Role of Economic Circumstances." *Oxford Bulletin of Economics and Statistics* 63 (2001): pp. 197-208.

[10] Brien, M., L. Lillard, and S. Stern. "Cohabitation, Marriage, and Divorce in a Model of Match Quality." *International Economic Review* 47 (2006): pp. 451-494.

[11] Browning, M., P.-A. Chiappori, and Y. Weiss. *Economics of the Family* (2014), Cambridge University Press.

[12] Bruze, G., M. Svarer and Y. Weiss. "The Dynamics of Marriage and Divorce." *Journal of Labor Economics* 33 (2015): pp. 123-170.

[13] Charles, K. K. and M. Stephens, Jr. "Job Displacement, Disability and Divorce." *Journal of Labor Economics* 22 (2004): pp. 489-522.

[14] Cherchye, L., T. Demuynck, B. De Rock and F. Vermeulen. "Household Consumption when the Marriage is Stable." *American Economic Review* 107 (2017a): pp. 1507-1534.

[15] Cherchye, L., T. Demuynck, B. De Rock and F. Vermeulen. "Stable Marriage with and without Transferable Utility: Nonparametric Testable Implications." (2017b). Mimeo.

[16] Cherchye, L., B. De Rock, A. Lewbel, and F. Vermeulen. "Sharing Rule Identification for General Collective Consumption Models." *Econometrica* 83 (2015): pp. 2001-2041.

[17] Cherchye, L., B. De Rock, and F. Vermeulen. "The Collective Model of Household Consumption: a Nonparametric Characterization." *Econometrica* 75 (2007): pp. 553-574.

[18] Cherchye, L., B. De Rock, and F. Vermeulen. "Opening the Black Box of Intra-Household Decision-Making: Theory and Nonparametric Empirical Tests of General Collective Consumption Models." *Journal of Political Economy* 117 (2009): pp. 1074-1104.

[19] Cherchye, L., B. De Rock, and F. Vermeulen. "The Revealed Preference Approach to Collective Consumption Behavior: Testing and Sharing Rule Recovery." *Review of Economic Studies* 78 (2011): pp. 176-198.

[20] Cherchye, L., B. De Rock and F. Vermeulen. "Married with Children: a Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure." *American Economic Review* 102 (2012): pp. 3377-3405.

[21] Chiappori, P.-A. "Rational Household labor Supply." *Econometrica* 56 (1988): pp. 63-89.

[22] Chiappori, P.A. "Collective labor Supply and Welfare." *Journal of Political Economy* 100 (1992): pp. 437-467.

[23] Chiappori, P. A. "Introducing Household Production in the Collective Model of Labor Supply." *Journal of Political Economy* 105 (1997): pp. 191-209.

[24] Chiappori, P. A. and I. Ekeland. "The Microeconomics of Efficient Group Behavior: Identification." *Econometrica* 77 (2009): pp. 763-799.

[25] Chiappori, P. A., B. Fortin and G. Lacroix. "Marriage Market, Divorce Legislation, and Household Labor Supply." *Journal of Political Economy* 110 (2002): pp. 37-72.

[26] Chiappori, P. A., S. Oreffice and C. Quintana-Domeque. "Bidimensional Matching with Heterogeneous Preferences: Education and Smoking in the Marriage Market." Forthcoming in *Journal of the European Economic Association* (2017).

[27] Chiappori, P. A., N. Radchenko and B. Salanie. "Divorce and the duality of marital payoff." *Review of Economics of the Household*, forthcoming (2018).

[28] Chong, A. and E. La Ferrara. "Television and Divorce: Evidence from Brazilian Novelas." *Journal of the European Economic Association* 7 (2009): pp. 458-468.

[29] Davison, J. "Tenacious Women: Clinging to Banja Household Production in the Face of Changing Gender Relations in Malawi." *Journal of Southern African Studies* 19 (1993): pp. 405-421.

[30] Davison, J.. *Gender, Lineage and Ethnicity in Southern Afric*a. Oxford: Westview Press, 1997.

[31] Diewert, E. "Afriat and Revealed Preference Theory." *Review of Economic Studies* 40 (1973): pp. 419-426.

[32] Doiron, D. and S. Mendolia. "Job Loss and Family Dissolution." *Journal of Population Economics* 25 (2012): pp. 367-398.

[33] Donni, O.. "Labor supply, home production, and welfare comparisons." *Journal of Public Economics* 92 (2008): pp. 1720-1737.

[34] Dunbar, G. R., A. Lewbel and K. Pendakur. "Children's Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi." *American Economic Review* 103 (2013): pp. 438-471.

[35] Dunbar, G. R., A. Lewbel and K. Pendakur. "Identification of Random Resource Shares in Collective Households Without Preference Similarity Restrictions." *Mimeo* (2017).

[36] Eliason, M. "Lost Jobs, Broken Marriages." *Journal of Population Economics* 25 (2012): pp. 1365-1397.

[37] Ellis, F., M. Kutengule and A. Nyasulu. "Livelihoods and Rural Poverty Reduction in Malawi." *World Development* 31 (2003): pp. 1495-1510.

[38] Farnham, M., L. Schmidt and P. Sevak. "House Prices and Marital Stability." *American Economic Review Papers & Proceedings* 101 (2011): pp. 615-619.

[39] Gale, L.S., and D. Shapley. "College admissions and the stability of marriage". *American Mathematical Monthly* 69 (1962): pp. 9-15.

[40] Greenwood, J., N. Guner, G. Kocharkov and C. Santos. "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review. Papers & Proceedings* 104 (2014): pp. 348-353.

[41] Hankins, S. and M. Hoekstra. "Lucky in Life, Unlucky in Love? The Effect of Random Income Shocks on Marriage and Divorce." *Journal of Human Resources* 46 (2011): pp. 403-426.

[42] Hitsch, G. J., A. Hortacsu and D. Ariely. "Matching and Sorting in Online Dating." *American Economic Review* 100 (2010): pp. 130-163.

[43] Jacquemet, N. and J.-M. Robin. “Assortative Matching and Search with Labor Supply and Home Production.” (2012). Mimeo.

[44] Jovanovic, B. “Job Matching and the Theory of Turnover.” *Journal of Political Economy* 87 (1979): pp. 972-990.

[45] Karlan, D., R. Osei, I. Osei-Akoto and C. Udry. “Agricultural Decisions after Relaxing Credit and Risk Constraints.” *Quarterly Journal of Economics* 129 (2014): pp. 597-652.

[46] Mazzocco, M. "Household Intertemporal Behavior: A Collective Characterization and a Test of Commitment." *The Review of Economic Studies* 74.3 (2007): pp. 857-895.

[47] Mwambene, L. “Divorce in Matrilineal Customary Law Marriage in Malawi: A Comparative Analysis with the Patrilineal Customary Law Marriage in South Africa.” LLM Thesis University of Western Cape (2005).

[48] Pollak, R.A. and M.L. Wachter. “The relevance of the household production function and its implications for the allocation of time.” *Journal of Political Economy* 83 (1975): pp. 255-277.

[49] Reniers, G. “Divorce and Remarriage in Rural Malawi.” *Demographic Research Special Collection* 1.6 (2003).

[50] Samuelson, P. “A Note on the Pure Theory of Consumers Behavior”. *Economica* 5 (1938): pp. 61-71.

[51] Takane, T. “Customary Land Tenure, Inheritance Rules and Smallholder Farmers in Malawi.” *Journal of Southern African Studies* 34 (2008): pp. 269-291.

[52] Udry, C. “Gender, Agricultural Production, and the Theory of the Household.” *Journal of Political Economy* 104 (1996): pp. 1010-1046.

[53] Varian, H. “The Nonparametric Approach to Demand Analysis.” *Econometrica* 50 (1982): pp. 945-972.

[54] Varian, H. “The Non-Parametric Approach to Production Analysis.” *Econometrica* 52 (1984): pp. 579-598.

[55] Voena, A. “Yours, Mine and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?” *American Economic Review* 105 (2015): pp. 2295-2332.

[56] Walther, S. “Moral Hazard in Marriage: The Use of Domestic labor as an Incentive Device.” *Review of Economics of the Household* 15 (2017): pp. 357-382.

[57] Walther, S. "Noncooperative Decision Making in the Household: Evidence from Malawi." *Journal of Development Economics*, Forthcoming (2018).

[58] Weiss, Y. and R. Willis. "Match Quality, New Information and Marital Dissolution." *Journal of labor Economics* 15.1.2 Essays in Honour of Yoram Ben-Porath (1997): pp. S293-S329.

[59] World Bank Development Economics Research Group and Malawi National Statistical Office. *Malawi Living Standards Measurement Study: 2010-2013 IHS3 Survey* (2015).

Appendix A: Proofs

Proof of Proposition 1

Necessity. To prove that the empirical conditions stated in Proposition 1 are necessary for the data set $\mathcal{D}$ to be consistent with a stable matching σ we apply the revealed preference argument that underlies Proposition 1 of Cherchye, Demuynck, De Rock and Vermeulen (2017a), but now adapted to our particular setting. Particularly, our conditions use information on (i) the bundles of goods consumed by individuals in their current match, (ii) the cost of these bundles in two alternative scenarios outside the observed match (i.e. as single (for the individual rationality requirement) and with some other potential partner (for the no blocking pair requirement)) and (iii) the available budget in these two counterfactual scenarios.

As explained in the main text, we assume that individuals are endowed with utility functions $U^a(l^a, q_a, Q)$ and $U^b(l^b, q_b, Q)$. For each matched couple $(a, \sigma(a))$, our data set $\mathcal{D}$ contains $l^a, l^{\sigma(a)}$ and $Q^{(a,\sigma(a))}$ and the aggregate private consumption $q^{(a,\sigma(a))}$. To reconstruct the individual consumption bundles, we have to consider all feasible specifications of $q_a^{(a,\sigma(a))}$ and $q_{\sigma(a)}^{(a,\sigma(a))}$ that satisfy $q_a^{(a,\sigma(a))} + q_{\sigma(a)}^{(a,\sigma(a))} = q^{(a,\sigma(a))}$ (i.e. condition (a)). For every observed match, this results in the individual consumption bundle $(l^a, q_a^{(a,\sigma(a))}, Q^{(a,\sigma(a))})$ for individual a and $(l^a, q_{\sigma(a)}^{(a,\sigma(a))}, Q^{(a,\sigma(a))})$ for individual $\sigma(a)$.

Next, in our labor supply model the price of an individual's leisure is the individual's wage, and the prices of the Hicksian private quantity $q^{(a,\sigma(a))}$ and the Hicksian public quantity $Q^{(a,\sigma(a))}$ are equal to one. We use this price information to compute the consumption cost of the within-marriage bundles in the two out-of-marriage scenarios. For the first scenario, if female a and male b would become single, they would have to bear the full cost of the public good to consume exactly the same quantity. When adding the cost of leisure and the private Hicksian quantities, this gives a total cost of $w^a l^a + q_a^{(a,\sigma(a))} + Q^{(a,\sigma(a))}$ for female a

and $w^b l^b + q_b^{(\sigma(b),b)} + Q^{(\sigma(b),b)}$ for male b.

For the second scenario, if the potentially blocking pair consisting of a and b would be matched, they would need the quantity $\max\{Q^{(a,\sigma(a))}, Q^{(\sigma(b),b)}\}$ of the public good to guarantee that both a and b consume at least the same amount as in their current match. Similarly to the first scenario, when adding the cost of leisure and the private Hicksian quantities, this yields a total cost of $\left(w^a l^a + w^b l^b\right) + \left(q_a^{(a,\sigma(a))} + q_b^{(\sigma(b),b)}\right) + \max\{Q^{(a,\sigma(a))}, Q^{(\sigma(b),b)}\}$ for the potentially blocking pair (a, b).

Restrictions (i) and (ii) in Proposition 1 compare these consumption costs with the available budget in the two counterfactual situations.[15] For each scenario, this available budget has three components. The first component is the potential labor income of each individual, i.e. $w^a T^a$ for female a and $w^b T^b$ for male b. The second component is the nonlabor income associated with the individuals' private land holdings. These private land holdings L^a and L^b are evaluated at the land prices $z^{(a,\sigma(a))}$ and $z^{(\sigma(b),b)}$, which generates the private land values $z^{(a,\sigma(a))} L^a$ for female a and $z^{(\sigma(b),b)} L^b$ for male b . These two first budget components are observed at the individual level, which means that we can assign these incomes to respectively a and b in the counterfactual scenarios. This assignability does not hold for the third budget component, which captures the remaining (non-assignable) nonlabor income, i.e. the sum of (1) nonlabor income $n^{(a,\sigma(a))}$, (2) the value $z^{(a,\sigma(a))} L^{(a,\sigma(a))}$) of the household's joint (non-assignable) land holdings and (3) the value $x^{(a,\sigma(a))}$ of other input used for agricultural production. To reconstruct the individual incomes of a and $\sigma(a)$, we have to consider all possible decompositions N^a and $N^{\sigma(a)}$ that satisfy the adding-up restriction $N^a + N^{\sigma(a)} = n^{(a,\sigma(a))} + x^{(a,\sigma(a))} + z^{(a,\sigma(a))} L^{(a,\sigma(a))}$ (similar to our treatment of the individual quantities $q_a^{(a,\sigma(a))}$ and $q_{\sigma(a)}^{(a,\sigma(a))}$).

As a final step, the individual rationality restrictions (i) in Proposition 1 state that a necessary condition for marital stability is that these individual budgets cannot strictly exceed the cost of the bundles consumed by the individuals in their current matches, which gives

$$N^a + z^{(a,\sigma(a))} L^a + w^a T^a \leq w^a l^a + q_a^{(a,\sigma(a))} + Q^{(a,\sigma(a))},$$
$$N^b + z^{(\sigma(b),b)} L^b + w^b T^b \leq w^b l^b + q_b^{(\sigma(b),b)} + Q^{(\sigma(b),b)}.$$

If these conditions were not met for some individual, then this individual would be better off by living alone for any possible specification of the individual (continuous, concave and

[15] We remark that our production assumption of profit maximization under constant returns to scale yields zero (maximum) profit. This implies that (1) total input value (used in our budget calculations) equals the value of the generated production output, and (2) there is no additional production profit (or loss) term to be included in the available consumption budgets.

monotonically increasing) utility functions. For example, the individual as a single could compose a consumption bundle with strictly more of each consumed good.

A directly analogous argument holds for the no blocking pair restrictions (ii) in Proposition 1. When evaluating the potentially blocking pair (a, b), we now compare the sum of the counterfactual budgets for female a and male b to the cost for a bundle guaranteeing at least the within-marriage consumption quantities for these two individuals. In this case, marital stability requires the inequality

$$\begin{aligned} &\left(N^a + z^{(a,\sigma(a))}L^a + w^a T^a\right) + \left(N^b + z^{(\sigma(b),b)}L^b + w^b T^b\right) \\ &\leq \left(w^a l^a + w^b l^b\right) + \left(q_a^{(a,\sigma(a))} + q_b^{(\sigma(b),b)}\right) + \max\{Q^{(a,\sigma(a))}, Q^{(\sigma(b),b)}\}. \end{aligned}$$

Sufficiency. Cherchye, Demuynck, De Rock and Vermeulen (2017b) introduced the Weak Axiom of Revealed Stable Matchings (WARSM) to define sufficient empirical conditions for a stable marriage allocation. Reformulating this WARSM for our setting gives exactly the conditions stated in Proposition 1. This shows that the data set $\mathcal{D}$ satisfies the empirical conditions in Proposition 1 if and only if it satisfies WARSM. Finally, Corollary 1 in Cherchye, Demuynck, De Rock and Vermeulen (2017b) states that the WARSM defines a sufficient condition for the data set $\mathcal{D}$ to be consistent with a stable matching σ as soon as all the inequalities in our condition (ii) are strict for the unmatched pairs (i.e. the pairs (a, b) with $b \neq \sigma(a)$).

Proof of Proposition 2

The result is an adaptation of Theorem 6 of Varian (1984) to our specific setting. In particular, we follow Chiappori (1997) by assuming profit maximization under constant returns to scale and exogenously given input and output prices. Let us start by assuming that the input prices of land (i.e. $z^{(a,\sigma(a))}$) and labor (i.e. w^a and $w^{\sigma(a)}$) are observed. Given that we assume a production technology with constant returns to scale, maximum attainable profit must equal zero. This defines the equality restriction (10) (for each observed match $(a, \sigma(a))$) in Proposition 2.

Next, profit maximizing behavior requires for every observed match $(a, \sigma(a))$ that, for the prices faced by $(a, \sigma(a))$, there does not exist a different input-output combination that yields higher profit. For a homogeneous production technology associated with a given marriage market, this yields the inequality restriction (11) in Proposition 2 for each combination of observed matches $(a, \sigma(a))$ and $(a', \sigma(a'))$. Intuitively, it says that $(a, \sigma(a))$ cannot attain a higher profit by adopting the input-output combination of $(a', \sigma(a'))$. Varian (1984, Theorem

6) has shown that consistency with these two requirements is a necessary and sufficient condition for the data to be consistent with profit maximizing behavior under a constant returns to scale production technology.

Finally, since we do not observe the input prices of land and labor, we simply need that there exists at least one possible specification of shadow land prices and wages that makes the data consistent with the profit maximization restrictions (10) and (11).

Appendix B: Data construction

All values used in our empirical application were converted to real terms using the spatial and temporal price index provided in the IHS. In some cases we recoded outliers, namely the top 1% of values, to be equal to the value at the 99th percentile.

Bounds on wages and land prices

Wages We calculated the median observed wage per hour of hired workers in the district, separately for males and females. Where there were insufficient observations, we used the regional median instead. The bounds were zero and two times this median.

Land price per acre For each plot of owned land, households were asked how much they could earn if they rented it out for one year. We regressed this value on plot characteristics: the size of the plot; the soil type of the plot; the soil quality of the plot; whether the plot is swamp or wetland; and how the household acquired the plot. We then used the predicted values of this regression to estimate the rental income for those plots where the reported rental income was missing. The rental income was summed for each household and divided by the total acres of land, giving an average rental income per acre for each household. We then obtained the median rental income per acre for each village and for each district. We used the median rental income per acre for the village where there were at least seven observations per village; where there were fewer, we used the median rental income per acre for the district. The bounds on the land price were zero and two times this median.

Production

Inputs We calculated the cost of inputs into production as the total of direct inputs, such as the costs of fertilizer, seeds and transport, the cost of indirect inputs, namely machinery, and the cost of hired labor. For machinery, we calculated the use value of each item by first calculating the remaining age of the item as twice the mean age of this item in the sample minus its current age, with a minimum of two years. The annual consumption stream from each item was the amount of money the item could be sold for, if sold today, divided by the remaining age of the item. The cost of hired labor was calculated as the number of days this labor was used times the average daily wage for these laborers, as reported by the household. The survey distinguished between male, female and child laborers, providing a more accurate measure of the total cost. Free labor was also valued at these rates and included as a costly input.

Revenue The revenue was calculated as the sum of all crop sales during the rainy and dry seasons and the value of all own agricultural production that was consumed by the household. The latter value originates from the survey itself, where households were asked how much of each consumed food they had grown themselves. This was then valued at local prices by the World Bank Living Standards Measurement Study team.[16]

Consumption

Consumption was split into four categories: public consumption; private non-assignable consumption; private consumption of the man and private consumption of the woman.

Public consumption This included expenditure on children's education and health, expenditure on the education and health of other household members (not the husband or wife), expenditure on children's clothing, expenditure on durables (which was calculated as a use value or consumption stream, using the same method described for machinery above), expenditure on public nondurables (such as candles, light bulbs and books), expenditure on rent and expenditure on public bills (such as firewood and the landline telephone).

Private non-assignable consumption The largest component of private non-assignable consumption was food, consisting of food purchased, the value of food from own production and the value of food received as a gift. This category also included private bills (such as the mobile telephone) and private nondurables (such as cigarettes, tickets for public transport, soap and stationery items).

Private consumption of the man and woman This consisted of the health, education and clothing expenses of the man or woman.

[16] Many thanks to Talip Kilic for sharing his Stata code that allowed us to separately identify consumption from own production and consumption from purchases.

Time

The model requires two time variables: agricultural labor and leisure.

Agricultural labor Agricultural labor was calculated as the total number of hours of agricultural work on the household's plots in the rainy and dry seasons of the past year, reported by the husband or wife. Where certain information was missing, such as the individual reported the number of days worked but not the number of hours per day, we used the village median for this information, where there were at least seven observations in the village. Otherwise, we used the district median.

Leisure In order to calculate leisure hours, we first required a measure of total available hours. As reported working hours are fairly low, leading to likely overestimates of true leisure time, we calculated total time available as the number of hours worked by the hardest working man or woman in the sample in the past year. This included both agricultural and wage labor and resulted in a value of 6120 hours. We assumed that this hardest worker works full-time and has zero leisure. We then calculated leisure for each individual as 6120 minus the annual hours of agricultural and wage labor of each individual.

Landholdings

In order to accurately measure the land income of individuals on divorce, we required exact information on the amount of land owned by each spouse. We defined land to be owned if it was inherited, granted by local leaders, part of a bride price, purchased with a title or purchased without a title. Land that was owned either solely by the spouse or owned by the spouse jointly with someone outside the household was assumed to accrue to that spouse on divorce. Land not owned by either spouse was assumed to disappear after divorce, while land owned jointly by the spouses was allowed to be endogenously split in the simulations.

Covariates in regressions

Here we explain how the covariates in the regressions were defined. All covariates from the data are from the 2010 wave. The 2013 wave was only used to see whether the couple had divorced.

 # Children This is the number of own or adopted children living in the household.

Age of man/woman This is the age of the man or woman in 2010.

Head education level This is a series of dummy variables that define the highest education level of the head, which ranges from no education to tertiary education.

Same age This is a dummy variable that equals one if the spouses have the same age ± 2 years, and equals zero otherwise.

Same education This is a dummy variable that equals one if the spouses have the same education level, and equals zero otherwise.

N Churches This is the number of churches in the marriage market, as reported by village informants.

Distance to road, urban centre This is the average distance to the nearest road or nearest urban centre (Lilongwe, Zomba or Blantyre) in kilometers, in the marriage market.

Sex ratio This is the ratio of men to women at the village level in the IHS sample, calculated based on the heads of household. Single-headed households count as one male or one female, while married households count as one male and one female.

Land This is the total number of acres of land owned by the household.

N Households in marriage market This is the total number of households in that particular household's marriage market.

Public/private share of consumption This is the share of public or private consumption in total consumption.

Nonlabor income (NLI) This is an output of the structural model and is the difference between total consumption and other inputs on the one hand and labor and land income on the other hand .

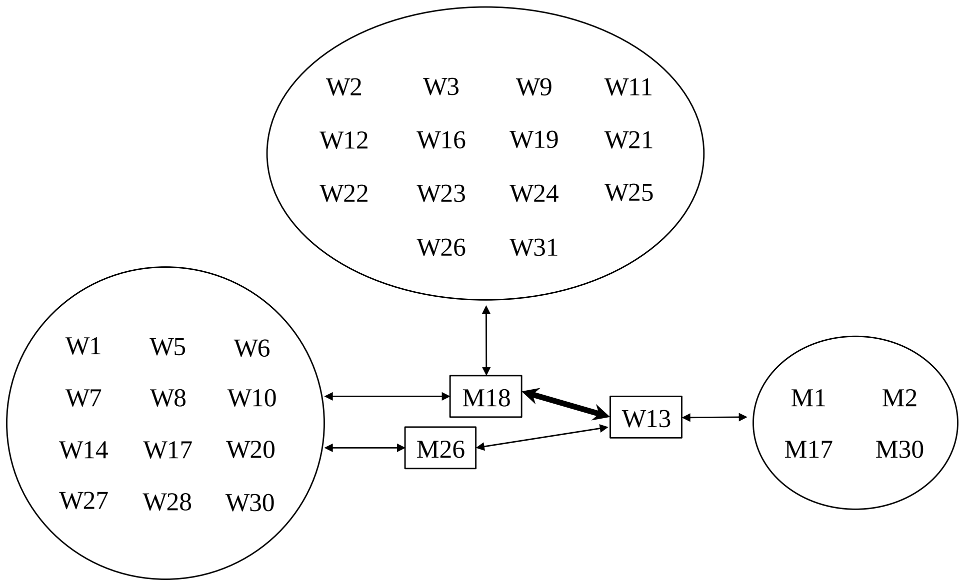
Land income This is an output of the structural model and gives the total number of acres of land owned by the spouse multiplied by the shadow price of land. It measures the annual rental yield on the land.

Wage This is an output of the structural model and gives the hourly shadow wage of agricultural labor of the husband or wife.

Appendix C: Example of instability network

In Figure 1 we illustrate the instability network of one particular cluster. Women are indexed Wi and men are indexed Mi, and we only display men and women who have blocking pairs in the cluster. Arrows depict these blocking pairs. In this cluster, $M18$, $M26$ and $W13$ are popular. $M18$ has a blocking pair with 27 women, meaning that he could be better off with any of these women than in his marriage, and each of these women would be better off with him. Similarly, $M26$ has a blocking pair with 13 women. $W13$ is the only woman with more than two blocking pairs: she has six. She can form a profitable blocking pair with $M1$, $M2$, $M17$ and $M30$, in addition to $M18$ and $M26$. However, she is best off with $M18$ (measured by the associated BP index). Similarly, $M18$ is best off with $W13$. The thick arrow depicts the fact that these two individuals are each others' favorite blocking pair: hence, they would both be best off divorcing their partners and marrying each other. The instability in this cluster is driven by $M18$, $M26$ and $W13$: if these three individuals were removed from the cluster, all marriages would be stable. The most likely explanation for the fact that these individuals have a large number of blocking pairs is that they are highly productive.

Figure 1: The instability network of one marriage market in our dataset. Arrows indicate profitable remarriage options; the thick arrow denotes the mutually best outside option (i.e. they are each others' BPmax).



Appendix D: Instability Indices and Household and Marriage Market Characteristics

Table 11: OLS regressions of instability indices in 2010 on household and marriage market characteristics in 2010

	(1) BPmax (w)	(2) BPavg (w)	(3) BPmax (m)	(4) BPavg (m)	(5) IR (m)
Primary edu.	-0.574***	-0.028***	-0.242	-0.149	-0.004*
	(0.219)	(0.010)	(0.181)	(0.092)	(0.002)
Secondary edu.	-0.648***	-0.027**	0.088	0.018	-0.004
	(0.250)	(0.012)	(0.186)	(0.094)	(0.003)
Tertiary edu.	-2.059***	-0.069***	-0.083	0.156	-0.003
	(0.532)	(0.025)	(0.449)	(0.328)	(0.008)
Age (m)	-0.003	-0.000	0.014*	0.007	0.000
	(0.010)	(0.000)	(0.007)	(0.004)	(0.000)
Age (w)	0.001	-0.000	-0.016**	-0.005	-0.000
	(0.011)	(0.000)	(0.008)	(0.005)	(0.000)
N Children	-0.095***	-0.004***	-0.053***	-0.021*	-0.001***
	(0.031)	(0.001)	(0.017)	(0.012)	(0.000)
Same age	-0.025	0.002	0.105	0.076	0.003
	(0.202)	(0.009)	(0.131)	(0.079)	(0.002)
Same edu.	-0.096	-0.011	-0.179	-0.136	-0.003
	(0.206)	(0.010)	(0.168)	(0.087)	(0.002)
MM Dist. to road	-0.042**	-0.002***	-0.025***	-0.005**	-0.000***
	(0.018)	(0.001)	(0.005)	(0.002)	(0.000)
MM Dist. to urban centre	-0.002*	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
MM N Churches	0.022	0.001	0.003	0.002	0.000
	(0.022)	(0.001)	(0.008)	(0.003)	(0.000)
MM N Households	0.045***	-0.001*	0.030***	-0.001	0.001***
	(0.013)	(0.001)	(0.007)	(0.002)	(0.000)
N	10545	10545	10545	10545	10545
R^2	0.076	0.025	0.056	0.005	0.104

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions.

Appendix E: Further results

This appendix displays further results tables. Table 12 shows logit regressions of divorcing between 2010 and 2013 on instability indices and control variables, parallel to the results in Table 7, with marginal effects evaluated at means. The estimated coefficients are very similar to those in the main table, although the coefficient on *BPmax* is not statistically significant. Next, Table 13 replaces *BPmax* in Table 7 in the main text with the 95th percentile of an individual's distribution of blocking pairs (including zeros). The results are, again, similar. Tables 14, 15 and 16 estimate the impact of the instability indices on remarriage versus being single. In particular, Table 14 estimates the effect of the average instability indices in an OLS regression, while Table 15 and 16 show relative risk ratios in multinomial logit regressions of the maximum and average indices, respectively. The results are generally consistent with the main text, although we find that *BPavg* significantly predicts divorcing and being single, in addition to remarriage, for the man. Further, the impact of the *IR* index, although insignificant, is very large in the logit model, and in the single outcomes in the multinomial models. This is indicative of insufficient variation in the IR index among divorced couples. Finally, Table 17 replicates the heterogeneity analysis in Table 10, but replacing the maximum indices with average indices. The results are similar, with significantly estimated negative effects of the number of children on the relationship between the *BP* indices and divorce. Put differently, the estimated positive effect of the *BP* indices on divorce is decreasing in the number of children that a household has in 2010.

Table 12: Logit regressions of divorce between 2010-2013 on instability indices in 2010 and other control variables

	(1)	(2)
	Divorced in 2013	
BPavg (woman)	0.017	
	(0.011)	
BPavg (man)	-0.001	
	(0.044)	
IR (man)	1.139	0.204
	(3.685)	(2.129)
BPmax (woman)		0.597**
		(0.249)
BPmax (man)		0.015
		(0.029)
N	5462	5462

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports Logit regressions. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head, the number of children the household had in 2010, and dummy variables indicating whether the couple are within two years of age of each other, and whether they have the same level of education.

Table 13: OLS regressions of divorce between 2010-2013 on 95th percentile instability indices in 2010 and other control variables

	(1)
	Divorced in 2013
BP95 (woman)	0.045***
	(0.016)
BP95 (man)	0.004
	(0.021)
IR (man)	0.424
	(1.650)
N	5847
R^2	0.128

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions. BP95 is the value of the 95th percentile of the individual's blocking pairs. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head, the number of children the household had in 2010, and dummy variables indicating whether the couple are within two years of age of each other, and whether they have the same level of education.

Table 14: OLS regressions of marital status in 2013 on average instability indices in 2010 and other control variables

	Marital status of man		Marital status of woman	
	(1) Remarried	*(2) Single*	*(3) Remarried*	*(4) Single*
BPavg (woman)	0.256***	0.288**	0.360***	0.089
	(0.085)	(0.128)	(0.131)	(0.061)
BPavg (man)	0.006	0.015	0.014	0.001
	(0.010)	(0.021)	(0.020)	(0.009)
IR (man)	-0.079	0.132	0.183	0.216
	(0.682)	(1.092)	(1.065)	(0.505)
N	1380	1380	1347	1347
R^2	0.113	0.109	0.123	0.093

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head, the number of children the household had in 2010 and dummy variables indicating whether the couple are within two years of age of each other, and whether they have the same level of education.

Table 15: Multinomial logit regressions of marital status in 2013 on instability indices in 2010 and other control variables

	(1) - Marital status of man		(2) - Marital status of woman	
	Remarried	*Single*	*Remarried*	*Single*
BPmax (woman)	1.124***	1.049	1.073**	1.046
	(0.037)	(0.060)	(0.031)	(0.039)
BPmax (man)	1.124*	1.016	0.917	1.027
	(0.037)	(0.132)	(0.110)	(0.075)
IR (man)	0.000	34.621	0.129	548.97
	(0.002)	(353.30)	(0.768)	(2879.05)
N	1347		1380	

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports the odds ratios from multinomial logit regressions. All regressions include district fixed effects (there was insufficient variation in outcomes within marriage markets), the age of the husband and wife in 2010, fixed effects for the education of the household head, the number of children the household had in 2010 and dummy variables indicating whether the couple are within two years of age of each other, and whether they have the same level of education.

Table 16: Multinomial logit regressions of marital status in 2013 on instability indices in 2010 and other control variables

	(1) - Marital status of man		(2) - Marital status of woman	
	Remarried	*Single*	*Remarried*	*Single*
BPavg (woman)	3.560***	7.156	19.137 **	4.214**
	(2.022)	(6.542)	(13.533)	(6.659)
BPavg (man)	1.069**	1.116	1.269	1.105
	(0.152)	(0.111)	(0.127)	(0.203)
IR (man)	0.001	18.798	0.001	0.726
	(0.003)	(96.438)	(0.006)	(6.312)
N	1347		1380	

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports the odds ratios from multinomial logit regressions. All regressions include district fixed effects (there was insufficient variation in outcomes within marriage markets), the age of the husband and wife in 2010, fixed effects for the education of the household head, the number of children the household had in 2010 and dummy variables indicating whether the couple are within two years of age of each other, and whether they have the same level of education.

Table 17: OLS regressions of divorce between 2010-2013 on average instability indices in 2010 and their interactions with other variables

	Divorced in 2013		*Divorced in 2013*		*Divorced in 2013*
	(1)		(2)		(3)
BPavg (woman)	1.069***	*BPavg* (woman)	0.483***	*BPavg* (female)	-0.368
	(0.257)		(0.164)		(0.920)
Age**BPavg* (woman)	-0.019***	Same age**BPavg* (woman)	0.095	Sex ratio**BPavg* (female)	0.956
	(0.005)		(0.268)		(1.023)
BPavg (man)	0.060**	*BPavg* (man)	0.017	*BPavg* (man)	-0.002
	(0.030)		(0.025)		(0.055)
Age**BPavg* (man)	-0.001***	Same age**BPavg* (man)	-0.014	Sex ratio**BPavg* (man)	0.019
	(0.000)		(0.015)		(0.054)
N	1406		1406		1406

	(4)		(5)
BPavg (woman)	0.015	*BPavg* (woman)	0.581***
	(0.009)		(0.191)
Same educ**BPavg* (woman)	-0.087	# Children**BPavg* (woman)	-0.042
	(0.200)		(0.035)
BPavg (man)	-0.011	*BPavg* (man)	0.030
	(0.030)		(0.027)
Same educ**BPavg* (man)	0.019	# Children**BPavg* (man)	-0.005*
	(0.014)		(0.003)
N	1406		1406

Standard errors in parentheses. * denotes p-value<0.1, ** denotes p-value<0.05 and *** denotes p-value<0.01. This table reports OLS regressions. All regressions include marriage market fixed effects, the age of the husband and wife in 2010, fixed effects for the education of the household head and the number of children the household had in 2010, and dummy variables for whether the couple are within two years of age of each other, and whether they have the same level of education. They also include the variable that is being interacted with the indices (e.g. Sex ratio in regression (3)).