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

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

November 2016

Abstract

Do individuals divorce for economic reasons? Can we measure the attractiveness of new matches in the marriage market? We answer these questions using a structural model of the household and a rich panel dataset from Malawi. We propose a model of the household with consumption, production and revealed preference conditions for stability on the marriage market. We define marital instability in terms of the consumption gains to remarrying another individual in the same marriage market, and to being single. Based on our estimates of marital instability in the first wave of the data, we find that a 1 percentage point increase in the wife's estimated consumption gains from remarriage is significantly associated with a 0.6 percentage point increase in divorce probability in the next three years. In a multinomial model, higher values of consumption gains from remarriage raise the odds of subsequent divorce and remarriage but not of divorce and singlehood. These findings provide out-of-sample validation of the structural model and shed new light on the economic determinants of divorce.

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

JEL classification: D11, D12, D13, J12

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

*We thank Ian Crawford, James Fenske, Arthur Lewbel, Alessandra Voena and seminar participants at University College London, Oxford, Rice, Texas A&M, IIES Stockholm, Tilburg, Warwick, Essex and ESEM 2016 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).

‡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). 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.

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. Such a 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 the particular union as well as on the opportunities implied by the marriage market as a whole.

The gains to marriage do not only consist of companionship and the production and rearing of children. There are also considerable economic gains, such as the sharing of public goods or the division of labor within unions (see Browning, Chiappori and Weiss, 2014, for an extensive discussion). The economic gains to marriage play a crucial role in the recent model proposed by Cherchye, Demuynck, De Rock and Vermeulen (2016) to analyze the impact of the marriage market on the intrahousehold distribution of resources. In their model, the collective model (Chiappori, 1988, 1992) is combined with the assumption of a stable marriage market, a concept that is directly related to the ideas in Becker (1973, 1974) and Becker, Landes and Michael (1977). The model predicts that the more attractive the outside options of a spouse, the higher is his or her share of the household's 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, e.g., Blundell, Chiappori, Magnac and Meghir, 2007, Cherchye, De Rock and Vermeulen, 2012, and Cherchye, De Rock, Lewbel and Vermeulen, 2015).

In this paper, we focus on the estimation of the economic gains to marriage and divorce in Malawi, which is one of the poorest countries in the world. Given the importance of agricultural production in this setting, we extend the theoretical model in Cherchye, Demuynck, De Rock and Vermeulen (2016) to include production. Appealingly, we can do this while allowing for both heterogeneous individual preferences and household production technologies. We estimate this model on panel data drawn from the Malawi Integrated Household Survey (IHS), and we test whether model-based measures of marital instability predict future divorces. As we discuss in more detail in Section 3.1, Malawi is an attractive setting for the estimation of this model because it has one of the highest divorce rates in Africa.[1] Reniers (2003) shows that in Malawi lifetime divorce probabilities are between 40% and 65%. He also shows that remarriage is almost universal: within two years of divorce, over 40% of women have remarried in the sample, and this figure reaches almost 90% after ten years. Thus, divorce and remarriage is a realistic outside option for married individuals.

Our model yields two structural indices for marriage instability: 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 (s)he remarried another individual in the same marriage market (the Blocking Pair (*BP)* index). In the

[1] Moreover, Dunbar, Lewbel and Pendakur (2013) also used the collective model to study the consumption behavior of Malawian households. These authors particularly focused on identifying resource shares of children. In the current study, we follow a similar collective approach to analyze marriage stability.

empirical analysis, we compute these instability indices for each married individual in the 2010 wave of our data. We then link these measures of instability to observed divorces between the 2010 and 2013 waves of the IHS. This sheds light on the importance of economic gains to marriage and divorce and how well our model is able to predict divorces and subsequent remarriages.

Our empirical results demonstrate that the wife's *BP* index significantly predicts divorce in the panel. A one-percentage-point increase in the wife's *BP* index, as a proportion of her household's income, raises the probability of divorce by 0.6 percentage points on average. This is a non-negligible effect, as the proportions of currently divorced and married individuals in the population suggest an annual divorce probability of 8.5%.[2] Therefore, we find that a model-based predictor of divorce correlates with out-of-sample realizations of divorce, hence validating the structural model. 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 effectively plays an important role in the gains to marriage and divorce. As an extension to these results, we also estimate a multinomial model that differentiates between individuals who divorce and remain single, and those who divorce and remarry. Interestingly, 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 idea that the *BP* index captures the attractiveness of remarriage.

Our paper makes two contributions to the literature. First, from a methodological point of view, it extends the model of Cherchye, Demuynck, De Rock and Vermeulen (2016) to account for domestic production in modeling households' economic behavior. This is particularly relevant for households in developing countries, for which agricultural production activities are prevalent. Interestingly, this also allows us to estimate shadow wages and land prices, which are often missing or suffer from measurement error in empirical applications. A distinguishing feature of our method is that it belongs to a revealed preference tradition that is free of any parametric assumptions, and thus obtains robust conclusions. See Samuelson (1938), Afriat (1967), Diewert (1973) and Varian (1982) for early contributions on revealed preference analysis of household consumption behavior. More recently, Cherchye, De Rock and Vermeulen (2007, 2009, 2011) have extended this seminal work towards the analysis of collective households.

Second, our empirical application offers a novel perspective on the growing literature on the economic drivers of divorce. This literature has focused on the relationship between economic shocks and divorce. For example, Weiss and Willis (1997) find that shocks to husbands' earning capacity reduce the probability of divorce, while shocks to wives' earning capacity increase it. In a similar vein, Boheim and Ermisch (2001) find that positive economic surprises among British households reduce the probability of union dissolution. Charles and Stephens (2004) compare the effect of spousal job loss and spousal disability, and find that job loss raises the probability of divorce but disability does not. Other papers examine the effect of lottery winnings on divorce (Hankins

[2] Assuming individuals change from marriage to divorce states according to a Markov process, using the proportions of individuals currently married and divorced in the IHS and using the remarriage probabilities in Reniers (2003), implies an annual divorce probability of 8.5%.

and Hoekstra, 2011) and the relationship between house prices and divorce (Farnham, Schmidt and Sevak, 2011). An alternative way to model observed patterns of cohabitation, marriage and divorce is proposed by Brien, Lillard and Stern (2006), who model match quality as an experience good, à la Jovanovic's (1979) labor market matching model. Overall, the evidence suggests that there is an important economic component to divorce, although these studies do not consider the potential for remarriage, and except for Brien, Lillard and Stern (2006), do not implement a structural measure of the economic gains to divorce. Jacquemet and Robin (2012) consider a search-matching model of the marriage market where individuals match on their wages and match quality, but divorce is exogenous. Our findings on the role of outside opportunities in triggering divorce complement their model: individuals match for financial reasons, but a high match quality can compensate for financial "mismatch". However, when match quality erodes, individuals search for a better financial match, and so divorce when there are more financially attractive individuals available in their marriage market.

As explained above, our structural approach combines the collective consumption model with the assumption of marriage stability, which allows us to explicitly incorporate the importance of intrahousehold sharing and options on the marriage market in the analysis of marriage and divorce decisions. The rest of this paper unfolds as follows. Section 2 introduces our revealed preference methodology for analyzing stability of marriage. Here, we also define our IR and BP indices for marriage stability. In Section 3 we provide further motivation for our empirical research question by explaining the specific setting of Malawi. We also present our sample of households and the construction of the marriage markets used in our empirical application. Section 4 presents some summary statistics on the main outcomes of our structural model. These results will motivate our regression analyses in Section 5, in which we focus on the relationship between the economic gains to divorce (captured by our IR and BP indices) and divorce and remarriage probabilities. Section 6 concludes.

2 Consumption, production and marriage stability

Our method for measuring instability of marriage builds on a recent paper by Cherchye, Demuynck, De Rock and Vermeulen (2016). These authors defined a revealed preference characterization of household consumption under stable marriage to analyze the intrahousehold allocation of resources. A novel feature of our analysis is that we integrate agricultural production in this revealed preference framework, thus linking productivity to marriage decisions. As we will explain further on (in Section 3.1), agricultural production is an important dimension of household decision behavior in Malawi. 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 convenient in view of our empirical application, because fully reliable wage and land price information is not available for the Malawian households that we study. As this last feature is often characteristic to data sets on developing countries, this also indicates the usefulness of our model in other application settings.

2.1 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 $\sigma : A \cup B \to A \cup B$. This function satisfies, for all $a \in A$ and $b \in B$,

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

In words, the function σ assigns to every female or male a partner of the other gender (i.e. $\sigma(a) = b$ and $\sigma(b) = a$). In this methodological section we will assume that $|A| = |B|$, which means that all individuals are matched. In contrast, in the empirical part of this paper we will account for the possibility that $|A| \neq |B|$ and that a married individual may consider remarrying a single of the other gender. Actually, if we do not include rationalizability conditions for singles' behavior, it is relatively straightforward to formally include this possibility in the models below.[3] However, unless there is a shortage on only one side of the marriage market, rationalizing the behavior of singles requires an explicit model for frictions on the marriage market and/or marriage costs. To focus our discussion, we abstract from such extensions in the current study.

Each individual i is assumed to spend her or his 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 $h^i \in \mathbb{R}_+$). The individual time budget constraint thus equals

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

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 input ($x \in \mathbb{R}_+$; e.g. fertilizer). We distinguish between land of the female ($L^a \in \mathbb{R}_+$), land of the male ($L^b \in \mathbb{R}_+$) and joint "household" land ($L^{(a,b)} \in \mathbb{R}_+$), i.e.

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

For a given match (a, b), we assume a common price for the three land types, i.e. the price of L^a, L^b and $L^{(a,b)}$ is given by $z^{(a,b)} \in \mathbb{R}_{++}$. 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

[3] Specifically, some of the variables in Propositions 1 and 2 (individual quantities, personalized prices, share of non-labor 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.

an agricultural production function $F\left(h^a, h^{\sigma(a)}, L, x\right)$. 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 at unity. The household is further associated with a 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 and $q_b \in \mathbb{R}_+$) and the household's expenditures on a public good (denoted by $Q \in \mathbb{R}_+$). Examples of private goods are expenditures on food and clothing, while an example of a public good is expenditure on children.

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.

2.2 Marriage stability: theoretical characterization

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

First, at the consumption level, we adopt the collective approach of Chiappori (1988, 1992) 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

$$\max_{l^a, l^{\sigma(a)}, q_a, q_{\sigma(a)}, Q} U^a\left(l^a, q_a, Q\right) + \mu U^{\sigma(a)}\left(l^{\sigma(a)}, q_{\sigma(a)}, Q\right) \tag{1}$$

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

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

where μ represents the Pareto weight of male $\sigma(a)$ relative to female a, and $N = n + x + zL$. We note that the Pareto weights are in general not constant. They will vary with factors such as wages. Next, we remark that N contains nonlabor income n as well as the rental value of land (quantities L^a, L^b and $L^{(a,b)}$, which are evaluated at the prices $z^{(a,b)}$ for each match (a, b)) and the cost for the other input x. In terms of the household's budget equation in (1), note that the total private consumption on the left-hand side of the budget equation contains both market consumption and agricultural output. Therefore we need to add the cost needed to produce this output to the right-hand side as well. These expenditures are, of course, equal to the expenses on the inputs needed to produce the specific output. The optimal composition of the inputs and the size of the agricultural output is determined next.

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

$$\max_{h^a, h^{\sigma(a)}, L, x} y - w^a h^a - w^{\sigma(a)} h^{\sigma(a)} - zL - x \tag{2}$$

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

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

Third, we assume that the marriage market is stable. Using the definition of Gale and Shapley (1962), marriage stability imposes that marriage matchings satisfy the conditions of Individual Rationality and No Blocking Pair. To formalize the notion of Individual Rationality, we let U_H^a and U_H^b represent female a's and male b's utility in their given marriage. Then, Individual Rationality requires

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

where U_S^a and U_S^b denote the female's and male's maximum attainable utilities as singles, respectively. Intuitively, Individual Rationality imposes that no female or male wants to exit the marriage and become single.

Next, to formalize the condition of No Blocking Pair, we let $U_{P_{(a,b)}}^a$ and $U_{P_{(a,b)}}^b$ 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

$$U_{P_{(a,b)}}^i > U_H^i \text{ implies } U_H^{i'} > U_{P_{(a,b)}}^{i'} \text{ for } i, i' \in \{a, b\}, i \neq i'. \tag{4}$$

In words, a marriage allocation has no blocking pairs if no female a and male b are both better off, with at least one individual strictly better off, by remarrying each other instead of staying with their current partner.

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 efficient.

2.3 Marriage stability: empirical conditions

For a given marriage market, let us for the moment assume that the data set $\mathcal{D}$ contains the following information:

- 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 $(a, \sigma(a))$.

As is standardly assumed for household data, 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. The individuals' private quantities will be treated as unknowns in our empirical conditions for marriage stability.[4] Next, in what follows we will assume that wages and land prices remain the same when individuals exit marriage (and become single or remarry), i.e. divorce has no productivity effects. The assumption that prices and wages are perfectly observed will be relaxed later on (see Section 2.4).

Characterizing stable marriage As explained in Section 2.2, 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, based on Cherchye, Demuynck, De Rock and Vermeulen (2016, Proposition 1), we obtain the following characterization of a stable marriage matching.[5]

Proposition 1 *The data set $\mathcal{D}$ is consistent with a* ***stable matching*** *σ only if there exist,*

a. for each matched pair $(a, \sigma(a))$ $(a \in A)$, individual quantities $q_a^{(a,\sigma(a))}$, $q_{\sigma(a)}^{(a,\sigma(a))} \in \mathbb{R}_+$ and nonlabor incomes N^a, $N^{\sigma(a)} \in \mathbb{R}_+$ that satisfy

$$q_a^{(a,\sigma(a))} + q_{\sigma(a)}^{(a,\sigma(a))} = q^{(a,\sigma(a))},$$
$$N^a + N^{\sigma(a)} = n^{(a,\sigma(a))} + x^{(a,\sigma(a))} + z^{(a,\sigma(a))} L^{(a,\sigma(a))},$$

[4] 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 A). Such information is easy to include in the linear conditions in Proposition 1. Basically, 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 will not explicitly consider this refinement here.

[5] After suitably adapting the notation of Cherchye, Demuynck, De Rock and Vermeulen (2016), the proof of Proposition 1 proceeds similarly as the proof of these authors' Proposition 1. Given this analogy, and for compactness, we do not explicitly include a formal proof in the current paper. Evidently, the proof can be obtained upon request.

b. and, for each pair (a, b) $(a \in A,\ b \in B)$, personalized prices $P_a^{(a,b)}$, $P_b^{(a,b)} \in \mathbb{R}_+$ that satisfy

$$P_a^{(a,b)} + P_b^{(a,b)} = 1,$$

such that the following constraints are met:

i. individual rationality restrictions for all females $a \in A$ and males $b \in B$, i.e.

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

ii. no blocking pair restrictions for all $a \in A$ and $b \in B$, i.e.

$$\begin{aligned} &\left(N^a + z^{(a,\sigma(a))}L^a + w^a T^a\right) + \left(N^b + z^{(\sigma(b),b)}L^b + w^b T^b\right) \\ &\leq \left(w^a l^a + w^b l^b\right) + \left(q_a^{(a,\sigma(a))} + q_b^{(\sigma(b),b)}\right) \\ &\quad + \left(P_a^{(a,b)} Q^{(a,\sigma(a))} + P_b^{(a,b)} Q^{(\sigma(b),b)}\right). \end{aligned} \tag{6}$$

Thus, 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 personalized prices $P_a^{(a,b)}$, $P_b^{(a,b)}$ that satisfy a set of constraints that are linear in these unknown quantities and prices. Therefore a convenient feature of the conditions in Proposition 1 is that they can be checked through simple linear programming, which means that they are easy to apply in practice. Also note that the observability of individual land constitutes a natural lower bound in conditions (5) and (6).

Conditions (a) and (b) of the proposition refer to feasibility restrictions with respect to the unknown individual quantities for matched pairs and with respect to personalized (Lindahl) prices for both matched and unmatched pairs. These conditions are associated with the assumption that households choose Pareto efficient intra-household allocations. Conditions (i) and (ii) can be given a "revealed preference" interpretation in terms of a stable marriage allocation. First, the inequalities (5) in condition (i) require, for each individual male and female, that the budget conditions 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 condition is not met, then at least one man or woman is better off (i.e. can attain a strictly better bundle) as a single, which means that the marriage allocation is not stable.

In a similar vein, the right hand side of the inequality (6) in condition (ii) gives the sum value of the bundles within marriage for female a (i.e. $w^a l^a + q_a^{(a,\sigma(a))} + P_a^{(a,b)} Q^{(a,\sigma(a))}$) and male b (i.e. $w^b l^b + q_b^{(\sigma(b),b)} + P_b^{(a,b)} Q^{(\sigma(b),b)}$), evaluated at the prices that pertain to the pair (a, b) (and using the personalized prices $P_a^{(a,b)}$ and $P_b^{(a,b)}$ to evaluate the public quantities). The inequality then requires that the pair's total income (i.e. $N^a + z^{(a,\sigma(a))}L^a + w^a T^a + N^b + z^{(\sigma(b),b)}L^b + w^b T^b$) must not exceed this sum value. Intuitively, if this condition is not met, then woman a and man b can allocate their income so that both of them are better off (with at least one strictly better off) than with their current matches $\sigma(a)$ and $\sigma(b)$, which makes (a, b) a blocking pair.[6]

Quantifying marriage instability An important focus of our empirical analysis will be 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: our Individual Rationality (IR) indices capture how much better off (in consumption terms) individuals would be as a single person, and our 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. become single or remarry 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 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 in (5) by

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

and the inequality (6) by

$$\begin{aligned}
\left(N^a + z^{(a,\sigma(a))}L^a + w^a T^a\right) &+ \left(N^b + z^{(\sigma(b),b)}L^b + w^b T^b\right) - s^{BP}_{a,b} \\
&\leq \left(w^a l^a + w^b l^b\right) + \left(q_a^{(a,\sigma(a))} + q_b^{(\sigma(b),b)}\right) \\
&\quad + \left(P_a^{(a,b)} Q^{(a,\sigma(a))} + P_b^{(a,b)} Q^{(\sigma(b),b)}\right),
\end{aligned} \tag{8}$$

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.

[6] 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.

In our application, we measure the degree of instability of our data set 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 pair 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 the household's income.

2.4 Shadow wages and land prices

So far we have assumed that prices are observed. If prices are not observed, we can use shadow prices. To do so, we can use the structural model that we defined in Section 2.2, which assumes profit efficient behavior under constant returns to scale (see (2)). In the spirit of Proposition 1, we use a characterization of profit efficiency that is nonparametric, which here means that it does not require an explicit specification of the production technology (represented by the function F).[7]

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 efficient production behavior. Specifically, we say that the data set $\mathcal{D}$ is consistent with shadow profit maximization if we can specify a production function F such that profit efficiency of the observed production behavior is supported by shadow prices. Adapting the notation in Varian (1984, Theorem 6) to our setting, we obtain the following characterization of productive efficient behavior.

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 $a' \in A$,

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

Basically, the conditions (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

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

(10)), which must exceed the profit for any household $(a', \sigma(a'))$ (with $a' \in A$) under the same prices (see (11)). 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 efficient 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

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

for $\underline{w}^a$, $\underline{w}^b$, $\underline{z}^{(a,\sigma(a))} \in \mathbb{R}_{++}$ and $\overline{w}^a$, $\overline{w}^b$, $\overline{z}^{(a,\sigma(a))} \in \mathbb{R}_{++}$ predefined lower and upper bounds. Appendix A explains how we define these bounds in our empirical application.

Our second extension pertains to the fact that the characterization in Proposition 2 implicitly assumes that different households are characterized by homogeneous production technologies. Clearly, in practice we need to account for unobserved technological heterogeneity across households, i.e. some households have access to less efficient production technologies than others. To account for this heterogeneity, 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 by heterogeneous technologies characterizing the different production processes.[8]

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

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

and the inequality restriction (11) by

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

Basically, the variables π^{a+}, π^{a-}, $\pi^{a,a'}$ account for deviations from the zero maximum profit that appears at the left hand side in the original conditions (10) and (11), i.e. they capture deviations from the assumption of productive efficiency under constant returns to scale with homogeneous household technologies.

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

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

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

subject to the constraints (a) and (b) in Proposition 1, the stability constraints (7) and (8) and the profit efficiency 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 simple linear programming.

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. As we use profit efficiency as our identifying assumption for the shadow wages and land prices, we set α very small. This can be interpreted in terms of a two-stage optimization process: in the first stage, we define shadow prices as the prices that correspond to minimal deviations from our profit inefficiency 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 $\sum_a s^{IR}_{a,\emptyset} + \sum_b s^{IR}_{\emptyset,b} + \sum_a \sum_b s^{BP}_{a,b}$).

3 Malawian households: setting, data and marriage markets

We start by sketching the specific context of Malawi. This will show that this country provides an interesting setting to investigate our question regarding the impact of economic determinants on marriage and divorce decisions. In a following step, we discuss our data selection and the construction of households' marriage markets.

3.1 The Malawi setting

Malawi is a poor country in Sub-Saharan Africa. The GDP per capita was \$226 in 2013 (World Bank). It ranks 174th out of 187 countries on the 2014 Human Development Index, with a life expectancy of 55.3 years at birth. This is partly due to the prevalence of HIV, which is one of the highest in the world, at 10% in 2014 (World Bank). Malawi ranked 129th out of 140 countries on the Gender Inequality Index, which measures inequality along three dimensions: reproductive health, empowerment and economic activity. It does better than surrounding countries partly because of the high proportion of female seats in parliament and the high female labor force participation rate. However, the proportion of females with secondary school education is low, at 10.4%.

According to Bignami-Van Assche et al (2011), around 90% of all employees work in the agricultural sector. Most of these are involved in smallholder production with land plots in the range of 0.2-3 hectares (Ellis, Kutengule and Nyasulu, 2003). The predominant crop grown is maize, and agricultural production involves the joint labor supply of husbands and wives (Telalagic Walther, 2015). Individuals' key assets and thus outside options, here defined as utility on divorce, are their landholdings and capacity for labor supply. Land is largely passed on through inheritance, often at the time of marriage (Walther, 2016). All this makes it clearly plausible that spouses divorce for economic reasons.

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 much higher figure than in other African countries (Reniers, 2003). Women are more likely to be divorced, separated or widowed than men (DHS Report, 2004), and marriage may be terminated either through a court decree or by the death of a spouse. This decree is relatively easy to obtain, as the spouse seeking divorce need only show that there is no love remaining in the marriage (Mwambene, 2005). 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, and thus a high turnover of divorces and remarriages, making it an appropriate setting for the model presented in Section 2, which assumes no frictions on the marriage market and remarriage or being single as realistic outside options.

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, authors' calculations). This allows us to be precise about defining the marriage markets within which divorced individuals can look for potential remarriage partners.

3.2 Household data and marriage markets

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.[9] 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. In the baseline survey, 768 communities were selected based on probability proportional to size, and within those 16 households were randomly sampled. The sample we use is restricted to rural households who report that they engage in agriculture.[10] We only include monogamous households where at least one spouse reports non-zero hours of agricultural labor in the past year. This produces a sample of 8624 married and single households. As explained above, we allow sin-

[9] Although data is available for two waves, we do not reestimate the structural model for the second wave because it is a smaller subsample of the first wave. The estimation of the instability indices and shadow prices relies crucially on information from other households in the marriage market, so that using a smaller sample would give results that are not comparable with those for the first wave.

[10] We use the survey weights provided in all our descriptive statistics and empirical analyses, and also take into account the fact that the primary sampling units are communities, so that clustering is at the community level, and that we are selecting a subpopulation from the original sample.

gles to form potential blocking pairs with married individuals. We obtain instability indices for 5924 married households, of which we observe 1404 households in the second wave. Appendix A discusses in more detail how the data for the estimation and empirical analysis were constructed.

A crucial component of our analysis is the definition of the marriage market, within which individuals can form potential blocking pairs. In Malawi, marriages tend to be local. In the IHS data set, 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 the marriage market. In particular, we use the GPS coordinates provided in the IHS data to construct clusters of two to three geographically close villages. We use the k-means algorithm in Matlab, which partitions the data into k clusters using the squared Euclidean distance. We set the number 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 makes it likely that individuals within a marriage market encounter each other. The more individuals there are in the marriage market, the more likely that there is a profitable new match. Thus, the size of the cluster can affect the values of instability. Therefore, we will control for the size of the cluster in our main analysis of divorce.

Table 1 describes the average age, education, number of children and consumption characteristics of our sample. We find that, on average, the household head is middle-aged and 76% of household heads have no education (education is measured by dummy variables equal to one if the household head's highest education is of that level, and zero otherwise). The average household has approximately three children and almost two acres of land.[11] 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.

4 Outcomes of the structural model

Table 2 summarizes some of the outputs of our structural model. We find that, on average, women have a significantly lower shadow wage than men, which is consistent with observed wages in Malawi. Women also have approximately one half of the land income of men, on average, which is partly driven by the fact that the average woman owns less land than the average man. However, women have significantly higher non-labor income than men. In our model, non-labor income captures the difference between consumption and agricultural income. High non-labor income is driven by low agricultural production, which in turn is driven by high hours of leisure in the sample. The high levels of non-labor income suggest under-reporting of agricultural labor. Although this leads to an underestimate of potential labor income on divorce, this underestimation is compensated for in non-labor income, so that the total potential incomes of spouses are not systematically affected by the under-reporting of agricultural labor.

[11] One acre equals 4046 m^2.

Table 1: Age, education and other characteristics of households in the sample

Variable	Mean
Age of head	40.39
	(0.22)
Head has no education	0.76
Head has primary education	0.10
Head has secondary education	0.12
Head has tertiary education	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	5924
Number of marriage markets	300

This table reports means and standard errors (between parentheses).

Table 2: Outputs of the structural model

	Men	Women
Wage	124.09	117.32
	(0.878)	(0.784)
Land income ('000s)	9.001	4.110
	(0.356)	(0.166)
Non-labor income ('000s)	137.82	206.30
	(3.187)	(2.771)
Number of observations	5924	

This table reports means and standard errors (between parentheses).

Table 3: Summary statistics of instability

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

This table reports means, standard errors (between parentheses) and % of non-zeroes.

Next, we describe the features of marriage instability in our sample. 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. Both indices are expressed as a percentage of the household's total income. Similarly, we also express the Individual Rationality (*IR*) indices as a percentage of the household's total income. Table 3 presents the summary statistics of these variables. Some interesting observations emerge. First, the estimated instability from potential blocking pairs shows that about 65% of women have a profitable match in their marriage market, while fewer than 17% of men have a profitable match. On the other hand, no woman in our sample would prefer to become single above staying married, while over 47% of men would do so. From the *BPmax* results, 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 gains. These results suggest that women have many unattractive potential matches and some very attractive potential matches, while men have mostly mediocre, somewhat attractive potential matches. We sharpen the intuition for our instability concepts in Appendix B, where we provide an example of the instability network of one cluster in our sample.

There are two ways of interpreting these findings. First, one can assume that the marriage market is frictionless, in the sense that any profitable opportunities are exploited. The model predicts that almost half of the men in our sample would like to be single and more than half of the women have profitable remarriage opportunities. However, given that the market is frictionless, the model must be omitting unobserved costs of being single for men and remarriage for women. For men, there may be an unobserved benefit to being married, such as the domestic labor of their wives. For women, there may be an unobserved cost of divorcing and remarrying, such as social stigma.

Second, one can assume that the marriage market has frictions and exploiting profitable opportunities takes time. In this case, the model predicts that many men in our sample will divorce to become single in the future, while many women will divorce and remarry. On the other hand, few women will choose to divorce and remain single, while few men will divorce and remarry. This is at odds with the prevalence of single-headed female households in Malawi. In what follows, we will shed more light on which of these two explanations is more likely by examining in more detail the changes in marital status between 2010 and 2013.

At this point, we note that the absence of domestic labor, which is currently included in leisure, in the model and data can explain the fact 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.

Table 4 shows the proportion of households that divorce between the 2010 an 2013 waves of the survey. While 1240 couples remain married, it is fair to say that the number of couples who split in this three-year period is relatively large, at 11.7% of the sample. 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. Of those women with known marital status in 2013, there is a fairly even share of single women and remarried women. On the other hand, most men divorce and remarry, with few remaining single. This is at odds with the assumption of frictions on the marriage market, which would imply that more men should become single rather than remarry between the two waves, and few women should divorce in order to be single.

Finally, Table 5 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, which suggests that these instability indices are capturing the returns to divorce. We will present a rigorous analysis of this relationship in Section 5.2. 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. This variable may be capturing marital duration, suggesting that couples who have been together longer are less likely to divorce (because they have weaker outside options, or because poor matches are dissolved early on).

Table 4: Marital status in the panel

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 5: Summary statistics by divorce status

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

This table reports means and standard errors (between parentheses).

5 Divorce and economic gains

We start by presenting some regression results that shed light on which variables are correlated with our stability indices. Subsequently, we further analyze the relation between our indices and observed divorces in our Malawian data set. We will conclude that our structural measures of marital stability are significantly related to observed divorces and remarriages. This indicates that divorce in Malawi is driven, at least partly, by economic motivations. Our empirical findings also provide out-of-sample validation of our structural model.

A preliminary remark pertains to our *BPmax* index. As explained above, this instability index has an intuitive interpretation as representing the gains of the most attractive remarriage option. However, it may be seen as a disadvantage that, by construction, its value increases monotonically with the number of observed remarriage possibilities. An index that does not suffer from this drawback, while having an intuition that is close to that of the *BPmax* index, is defined as the 95^{th} percentile of the distribution of the *BP* indices for a married individual. We have verified the robustness of our following results to using this 95^{th} percentile index. Reassuringly, our regressions using the 95^{th} percentile index yield the same qualitative conclusions as the ones for the maximum index. For compactness, we will not report these regression results here, but they are available upon request.

5.1 What drives instability?

In the first step of our empirical analysis, we examine how household characteristics can explain instability, in particular the *BPmax* and *BPavg* indices of men and women. We explore the effect of characteristics such as the head's education level, the consumption quintile, landholdings and the distance to the nearest road and urban area. The consumption quintiles are dummy variables that equal one if the household's per capita consumption is in that bracket, and zero otherwise; for example, the fifth quintile is a dummy variable equal to one if the household's per capita consumption is in the top 20% of the sample. As the dependent variables are censored below zero, we perform tobit regressions. We present the marginal effect of covariates at the means of these covariates on the censored variable. Table 6 presents these results.

We find that the more educated the household head, the lower are the wife's *BP* indices (i.e. her remarriage possibilities are less attractive). On the other hand, a secondary school educated household head leads to higher instability for the man, compared to no education. This pattern can be explained by the fact that most household heads are male, so that a highly educated man is a more attractive husband, both to his wife and to other women in the marriage market. The results also show that wealthier households are more stable: this is captured by landholdings and the consumption quintile. Women, in particular, appear to be more maritally stable when their household owns more land, while being in the top consumption quintile has an especially significant stabilizing effect, as compared to being in the bottom consumption quintile (the excluded group). These results are consistent with the descriptive statistics in Table 5, where households

Table 6: Explaining variation in instability: Tobit marginal effects

	(1)	(2)	(3)	(4)
	BPmax (woman)	*BPavg* (woman)	*BPmax* (man)	*BPavg* (man)
Head educ.: Primary	-0.216	-0.011	0.094	0.051
	(0.180)	(0.007)	(0.070)	(0.038)
Head educ.: Secondary	-0.470***	-0.019**	0.243***	0.143***
	(0.170)	(0.007)	(0.080)	(0.049)
Head educ.: Tertiary	-1.672***	-0.058***	0.580	0.393
	(0.490)	(0.022)	(0.565)	(0.349)
2nd Quintile (per capita cons.)	0.061	0.002	-0.154	-0.085
	(0.190)	(0.008)	(0.100)	(0.058)
3rd Quintile	-0.257	-0.009	-0.190**	-0.110**
	(0.206)	(0.010)	(0.092)	(0.051)
4th Quintile	-0.166	-0.008	-0.330***	-0.175***
	(0.217)	(0.009)	(0.092)	(0.052)
5th Quintile	-0.869***	-0.035***	-0.449***	-0.247***
	(0.206)	(0.009)	(0.092)	(0.053)
Land (acres)	-0.106**	-0.004***	0.000	-0.003
	(0.043)	(0.002)	(0.014)	(0.008)
HH Distance to urban centre, km	-0.004*	-0.000	0.000	-0.000
	(0.002)	(0.000)	(0.000)	(0.000)
HH Distance to road, km	-0.061***	-0.002**	-0.016***	-0.007***
	(0.023)	(0.001)	(0.005)	(0.003)
Number of observations	5924	5924	5924	5924

This table reports the results of a tobit regression, with marginal effects evaluated at the means of the covariates. District fixed effects and age of household head in all regressions. * indicates significance at the 10% level, ** at the 5% level and *** at the 1% level.

who divorce in the panel own less land and have lower consumption. Finally, we observe an effect of connectedness on stability: the closer households are to the nearest road, the more unstable they are; this is true for both spouses. A one kilometer increase in the household's distance to the nearest road reduces the wife's *BPmax* (i.e. maximum consumption gain from remarriage as a percentage of household income) by 0.06 percentage points and the husband's by 0.02 percentage points. This may be because being more connected to other places makes it easier for spouses to widen their marriage market.

5.2 Divorce

We now present the main empirical analysis. We examine whether our instability indices, estimated using 2010 data, can predict divorce in the following three years. We estimate a logit model of divorce between 2010 and 2013, with the *BP* indices of the spouses and the *IR* index of the husband as covariates. We do not include the wife's *IR* index because this measure does not not vary in our sample. In addition to the covariates in the regressions in Table 6, we also include variables to measure religiousness in the village (the number of churches and the number of mosques), an additional measure of connectivity (access to a telephone in the village) and measures of match quality (the education, age and religion difference of the spouses). Finally, we also include the size of the cluster as a control variable. The results are in Table 7.[12]

Interestingly, we find a significant relationship between the instability indices from our structural model and subsequent divorce. In regression (1), a one-percentage-point increase in the wife's maximum gain from remarriage, as a proportion of her household's income, raises the probability of divorce by 0.6 percentage points on average. This is a non-negligible effect, as the proportions of currently divorced and married individuals in the population suggest an annual divorce probability of approximately 8.5%. In regression (2), a one-percentage-point increase in the average remarriage gain for the wife, as a proportion of her household's income, raises divorce probability by 17.9 percentage points. Note that the impact of a percentage point change of the maximum and average gains from remarriage on the divorce probability are not directly comparable to each other since the base levels of the *BPmax* and *BPavg* are very different (see Table 3). The *BPavg* index of the husband also has a positive, significant effect on divorce probability, although the effect is small in magnitude compared to that of the wife's *BPavg* index. Next, we find that the *IR* index of the man has a slightly negative, but statistically significant, impact on divorce. This may seem surprising at first sight. However, as we will explain further on (when discussing Table 9), this negative effect of the male's *IR* index disappears if we condition on wages, non-labor incomes and land incomes and, therefore, it may be regarded as mainly capturing an income effect. Overall, these results suggest that our measures of instability are able to capture the gains to divorce, and that women in Malawi are more likely to divorce for economic reasons than men. These results validate our structural

[12] As the instability indices are generated variables, one could estimate bootstrapped standard errors. However, we have opted not to do this in the current analysis due to the computational load, particularly because we are using weights.

Table 7: A logit model of marital instability and divorce

	(1) Maximum *BP*s	(2) Average *BP*s
	Divorced in 2013	
BP (woman)	0.006***	0.179***
	(0.002)	(0.049)
BP (man)	0.004	0.016**
	(0.005)	(0.007)
IR (man)	-0.002	-0.006*
	(0.004)	(0.004)
# Children	-0.014**	-0.015**
	(0.006)	(0.006)
Age of head	-0.002**	-0.002**
	(0.001)	(0.001)
2nd Quintile	-0.098***	-0.099***
	(0.030)	(0.030)
3rd Quintile	-0.064**	-0.066**
	(0.027)	(0.027)
4th Quintile	-0.069**	-0.073**
	(0.031)	(0.031)
5th Quintile	-0.039	-0.046
	(0.037)	(0.036)
# Churches	0.001	-0.000
	(0.002)	(0.002)
# Mosques	-0.011**	-0.011**
	(0.005)	(0.005)
Access to phone	0.044*	0.051**
	(0.025)	(0.023)
Size of cluster	0.000	0.001
	(0.001)	(0.001)
Number of observations	1404	1404

This table reports the results of a logit model, with marginal effects evaluated at the means of the covariates. Both regressions also include the age, education and religion difference of the spouses, the education level of the household head, the HH distance to the nearest road and urban centre and district fixed effects. * indicates significance at the 10% level, ** at the 5% level and *** at the 1% level.

model.

Other covariates are also significantly related to divorce. The probability of divorce is decreasing in the number of children, with an additional child reducing divorce probability by 1.4 percentage points. This implies that an approximately 2.3 percentage point reduction in the wife's maximum gain from remarriage, as a percentage of income, reduces divorce probability as much as an additional child. The probability of divorce is falling in the age of the household head, which may be because couples who are together longer are better matched, or because the value of outside options on the marriage market falls with age. We also find that divorce probability is decreasing in the household's wealth, as captured by the per capita consumption quintile. This is consistent with the descriptive statistics in Section 4. In villages with many mosques, divorce is less likely, while divorce is more likely in villages with a telephone, again suggesting that connectedness plays an important role in household dissolution. Finally, the size of the cluster is not significantly related to divorce.

It is worth noting that the absence of domestic labor cannot explain the significant effect of the wife's instability index on divorce. As domestic labor is currently included in leisure, marriages appear to be more attractive than they actually are. Consider a woman who engages in a lot of domestic labor: she appears to be stable, but at the same time is 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. Finally, our results cannot be explained by polygamy. The inclusion of a dummy for the existence of polygamy in the village does not affect the significant 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 available on request).

5.3 Extensions

Multinomial model 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 on the marriage market, while the *IR* index measures the attractiveness of being single. Therefore, we should observe these associations in the data as well. In order to explore this, we estimate a multinomial logit model for the marital status of husbands and wives in 2013, distinguishing between remarriage and being single.[13] We retain the same right-hand side variables as in Table 7 and set remaining married as the base case.[14] The results are in Table 8, which reports relative risk ratios (exponentiated

[13]The multinomial logit model assumes Independence of Irrelevant Alternatives (IIA). A more general model is the nested logit, which allows for correlation between alternatives. We estimated a nested logit model for marital status with a reduced set of district dummy variables, as the model did not converge with the full set. In this restricted version of the nested logit model, the IIA assumption was not rejected. Therefore, we proceeded with the multinomial logit model.

[14]We only report results for the *BPmax* index to save space. However, using the *BPavg* index yields the same qualitative conclusions. The main difference is that women's *BPavg* predicts divorce and being single for women. This may be due to informational constraints: women divorce in the hope that they will remarry, but find that they

coefficients).[15]

The results are consistent with the premise that the *BP* indices measure the attractiveness of remarriage. In particular, a higher value of the husband's *BP* index is significantly associated with a higher probability that the husband divorces and remarries in the next three years, instead of remaining married. The effect of the wife's *BP* index is stronger and is significantly associated with the remarriage of both the husband and the wife. This result is consistent with the observation that in Malawi, men find it easy to remarry. A one percentage point increase in the wife's maximum remarriage gain, as a proportion of household income, raises the relative risk of both the husband and wife divorcing and remarrying, relative to remaining married, by a factor of 1.1. Neither the wife's nor the husband's *BP* index raises the odds of divorcing and being single, relative to staying married; this is encouraging, as blocking pairs relate specifically to potential remarriage partners and should not be related to individuals divorcing in order to be single. Therefore, the *BP* indices appear to capture the attractiveness of remarriage in particular. The husband's *IR* index does not affect the odds of either divorce status, relative to remaining married. This may be because the values of the *IR* indices are very low in general, with little variation; because few men are observed to be single in the data; or because there are unobserved benefits to being married.

Other significant effects persist from Table 7: the odds of a woman divorcing and remarrying are declining in the number of children, as are the odds that a man divorces and remains single. The number of mosques has a significant negative effect on the odds that the man divorces and remarries, as well as the odds that the woman divorces and remains single. The presence of a telephone in the village increases the odds of remarriage for both the husband and wife but, interestingly, has no effect on the odds that either spouse is single. This supports the idea that the telephone captures some of the ability of spouses to learn about their marriage market. Finally, a higher consumption quintile of the household is broadly associated with lower odds of divorce for the wife, consistent with the results in Table 7, which showed that wealthier households are less likely to divorce.

Can the effect of instability be explained by wages or income? The results so far suggest that individuals, and particularly women, in Malawi divorce for economic reasons and that our instability indices capture these economic reasons well. However, one might argue that the *BP* index does not capture the economic gains from divorce, but rather is entirely driven by one or more of its components from equation (8) in Section 2.3. In order to explore this possibility, we introduce these components as explanatory variables to regressions (1), (2) and (3) in Table 7. In particular, we include the average wage of the husband and wife, the difference between these wages, the average non-labor income of the spouses, its difference, as well as the average land income, its difference, and the log of total income. All of these variables are the product of the estimation discussed in Section 2.4. The results are in Table 9.

are unable to find a suitable partner. This echoes the heterogeneous effects depending on sex ratio that we find in the later part of this section. The regression results are available upon request.

[15] The sample size in these estimates is lower than in the previous table because we do not know the marital status of every divorced man and woman (see also the explanation in Section 3.2).

Table 8: Multinomial logit regressions on marital status

	(1) - Marital status of man		(2) - Marital status of woman	
	Remarried	*Single*	*Remarried*	*Single*
BPmax (woman)	1.134***	0.968	1.088**	1.033
	(0.036)	(0.062)	(0.039)	(0.039)
BPmax (man)	1.160*	1.004	0.927	1.025
	(0.091)	(0.168)	(0.116)	(0.065)
IR (man)	0.944	1.023	0.945	1.068
	(0.064)	(0.115)	(0.058)	(0.057)
# Children	0.961	0.628***	0.804**	0.855
	(0.090)	(0.063)	(0.072)	(0.096)
Age of head	0.970*	0.995	0.967**	0.978
	(0.016)	(0.022)	(0.015)	(0.014)
2nd Quintile	0.451	0.074**	0.325**	0.323**
	(0.221)	(0.094)	(0.144)	(0.178)
3rd Quintile	0.586	0.182**	0.324***	0.650
	(0.273)	(0.145)	(0.142)	(0.290)
4th Quintile	0.466	0.295*	0.503*	0.198***
	(0.257)	(0.217)	(0.209)	(0.105)
5th Quintile	1.498	0.011***	0.313*	0.494
	(0.770)	(0.018)	(0.190)	(0.312)
# Churches	1.001	0.964	0.990	1.033
	(0.022)	(0.046)	(0.031)	(0.036)
# Mosques	0.830***	0.857	0.949	0.883*
	(0.059)	(0.094)	(0.062)	0.061
Access to phone	1.908*	0.470	2.050**	1.505
	(0.716)	(0.368)	(0.616)	(0.591)
Size of cluster	1.005	0.989	1.003	1.024*
	(0.009)	(0.026)	(0.010)	0.014
Number of observations	1345		1378	

This table reports relative risk ratios (standard error) from a multinomial logit model. All regressions also include the age, education and religion difference of the spouses, the education level of the household head, the HH distance to the nearest road and urban centre and district fixed effects. District FEs for regression (2) are an aggregated version of those for regression (1), due to insufficient variation in the outcome in some districts. * indicates significance at the 10% level, ** at the 5% level and *** at the 1% level.

The key result is that the coefficients on the *BP* indices in this table are not significantly different from those in the regressions in Table 7. In other words, the effect of our instability index cannot be explained simply by differences in wages, non-labor income or land income of the spouses. These variables do seem to explain the negative impact of the *IR* index of the man, however, indicating that this index is mainly capturing an income effect. Thus, we can conclude that the *BP* indices are able to capture a more complex form of gains from divorce that likely includes intrahousehold sharing of consumption, which is an important determinant of the *BP* indices in the model. Finally, we find that the average wage and the non-labor income of the household have a significant decreasing effect on the probability of divorce, beyond any indirect effect through the *BP* indices. Presumably, this is because they increase the gains to the current marriage. The same conclusion does not hold for land income and total income, but note that the impact of these variables is rather small compared to all the other variables.

Heterogeneous effects We have shown that spouses' *BP* indices are significantly associated with divorce, that this result holds in a multinomial model and that it cannot be explained by the components of the *BP* indices from the structural model. This relationship may mask significant heterogeneity, and this is what we explore here. We focus on the most attractive remarriage possibility (*BPmax*) and we include some new variables in the logit model of Table 7, such as landholdings and the sex ratio (defined as the ratio of males over females in a given cluster), as well as explore heterogeneity in existing explanatory variables, namely age of the household head and number of children. Table 10 shows these results.[16]

In regression (1), we find that the significant effect of *BPmax* on divorce only holds when the household has positive landholdings, and is increasing in these landholdings. For a household that owns two acres of land, which are the average landholdings in our sample, an increase in the wife's *BPmax* of one percentage point increases the probability of divorce by 0.6 percentage points, which is comparable to the effect in Table 7. The result also makes intuitive sense: land provides security on divorce, so spouses without land may find it too economically risky to divorce.

Next, we consider the interaction between age and the *BPmax* indices. Regression (2) shows that the wife's *BPmax* index is still a significant predictor of divorce, but this effect is declining in the age of the household head. This may be for two reasons: first, age may be a proxy for marital duration, and assuming match quality is revealed over time, spouses may be less driven by economic incentives later on in marriage. Second, age may tell us something about the individual's outside options. An older divorcée may have lower chances on the remarriage market than a younger divorcée, all other things equal, so that she may be less likely to respond to attractive outside options.

In the baseline regressions in Table 7, children always reduce the probability of divorce. In

[16]We only give results for the *BPmax* index to save space. However, using the *BPavg* index yields the same qualitative conclusions. In fact, Sex ratio**BPavg* (man) in regression (4) has a negative and significant effect on divorce probability. This is a rational response since for men, an increase in the sex ratio represents fewer women relative to men. In this sense, the results reaffirm the fact that individuals' response to their BP index depends on the supply of potential marriage partners. The regression results are available upon request.

Table 9: Explaining the effect of instability with its individual components

	(1) Maximum BPs	(2) Average BPs
	Divorced in 2013	
BP (woman)	0.006***	0.161***
	(0.002)	(0.048)
BP (man)	0.002	0.013*
	(0.005)	(0.007)
IR (man)	0.001	-0.004
	(0.004)	(0.004)
Wage (avg.)	-0.005**	-0.006**
	(0.002)	(0.002)
Wage (diff.)	-0.001	-0.001
	(0.001)	(0.001)
NLI (avg., 000s)	-0.001**	-0.001**
	(0.000)	(0.000)
NLI (diff., 000s)	0.000	0.000
	(0.000)	(0.000)
Land income (avg., 000s)	-0.002	-0.002
	(0.001)	(0.001)
Land income (diff., 000s)	0.001*	0.001*
	(0.001)	(0.001)
Log(total income)	0.539**	0.574**
	(0.225)	(0.223)
Number of observations	1404	1404

This table reports the results of a logit regression, with marginal effects evaluated at the means of the covariates. Both models also include the same right-hand side variables as in Table 7. * indicates significance at the 10% level, ** at the 5% level and *** at the 1% level.

Table 10: Heterogeneous effects of blocking pairs

	(1)		(2)
	Divorced in 2013		*Divorced in 2013*
BPmax (woman)	0.001	*BPmax* (woman)	0.018**
	(0.003)		(0.007)
Land**BPmax* (woman)	0.003***	Age**BPmax* (woman)	-0.0003*
	(0.001)		(0.000)
BPmax (man)	0.007	*BPmax* (man)	0.024
	(0.008)		(0.019)
Land**BPmax* (man)	-0.003	Age**BPmax* (man)	-0.001
	(0.003)		(0.001)
Land		Age	-0.001
	(0.011)		(0.001)
Number of observations	1404	Number of observations	1404

	(3)		(4)
	Divorced in 2013		*Divorced in 2013*
BPmax (woman)	0.011***	*BPmax* (female)	-0.040***
	(0.004)		(0.015)
Children**BPmax* (woman)	-0.002	Sex ratio**BPmax* (female)	0.056***
	(0.001)		(0.018)
BPmax (man)	0.018*	*BPmax* (man)	0.019
	(0.009)		(0.020)
Children**BPmax* (man)	-0.007**	Sex ratio**BPmax* (man)	-0.015
	(0.003)		(0.022)
Children	-0.008	Sex ratio	-0.225
	(0.008)		(0.154)
Number of observations	1404	Number of observations	1404

This table reports the results of a logit model, with marginal effects evaluated at the means of the covariates. All regressions in this table include the same right-hand side variables as in Table 7. * indicates significance at the 10% level, ** at the 5% level and *** at the 1% level.

regression (3), we find that this is especially true for men: high remarriage gains for the husband are less likely to result in divorce if the couple has more children. This is also true for the wife, although the effect is not statistically significant. Finally, in regression (4) we examine the interaction between the sex ratio and the effect of blocking pairs on divorce. For a sex ratio equal to one, an increase in the wife's $BPmax$ index of one percentage point increases the probability of divorce by approximately 1.6 percentage points. The more men there are, relative to women, the stronger the 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.

6 Conclusion

We have defined structural measures of the gains from divorce and shown that they are significant predictors of future divorces. These measures are based on a collective model with consumption and agricultural production embedded in a marriage market. The key theoretical contribution is that we extend Cherchye, Demuynck, De Rock and Vermeulen (2016) to include agricultural production, allowing the estimation of shadow wages and land prices. The model yields marital stability conditions for each married individual. Using these conditions, we can 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).

We estimate our model on the 2010 wave of the Malawi Integrated Household Survey, and correlate our instability indices with divorce in the next three years of the panel data set. We find that a 1 percentage point increase in the wife's BP index as a proportion of her household income increases divorce probability by 0.6 percentage points in the next three years. This result is robust to using an average BP or most attractive BP index, as well as to the estimation of a more general multinomial model. We also show that this result cannot be explained by the wages or land income of the spouses, implying that intrahousehold sharing matters.

Our findings lead us to conclude that divorce in Malawi is driven, at least partly, by economic considerations of spouses, with consumption sharing within households as an important determinant. In addition, our empirical results validate the set-up of our theoretical model. More generally, they show the value-added of adopting a Beckerian approach that analyzes marriage decisions through the lens of a structural model of household decision making. Finally, as agricultural productivity is a key determinant of outside options in developing countries reliant on agriculture, our model is applicable to other contexts as well, where accurate data on wages or land prices may not be available.

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

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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

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

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

[12] Cherchye, L., T. Demuynck, B. De Rock and F. Vermeulen. "Household Consumption when the Marriage is Stable." *American Economic Review,* Forthcoming (2016).

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

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

[15] 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.

[16] 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.

[17] 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.

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

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

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

[21] 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.

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

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

[24] 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.

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

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

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

[28] 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.

[29] Kerr, R. B. "Informal Labor and Social Relations in Northern Malawi: The Theoretical Challenges and Implications of Ganyu labor for Food Security." *Rural Sociology* 70 (2005): pp. 167-187.

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

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

[32] 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).

[33] 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.

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

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

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

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

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

[39] Walther, S. "Moral Hazard in Marriage: The Use of Domestic labor as an Incentive Device." *Review of Economics of the Household,* Forthcoming (2016).

[40] Walther, S. "Noncooperative Decision Making in the Household: Evidence from Malawi." (2016). Mimeo.

[41] 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.

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

Appendix A: 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. This was because in those cases the top 1% of values were not sensible, given the context of the data.

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.[17]

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

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.

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.

Consumption quintiles This splits the entire sample into five sections, in terms of per capita consumption. The first quintile is a dummy variable that equals one if the household is in the bottom 20% of per capita consumption in the full IHS sample and zero otherwise, while the fifth quintile is a similar definition for households in the top 20%. Other quintiles are defined analogously.

Age of head This is the age of the household head; the identity of the head was self-reported in the data.

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.

Diploma difference This is the husband's highest achieved diploma minus the wife's highest achieved diploma. The highest achieved diploma ranges from 1 (none) to a 7 (post-graduate degree). The difference ranges from -3 to 6 in our sample.

Age difference This is the husband's age minus the wife's age.

Religion difference This is a dummy variable that equals one if the spouses are of a different religious denomination, and zero otherwise.

 # Churches/ #Mosques This is the number of churches/mosques in the village, as reported by village informants.

Distance to road, urban centre This is the household's distance to the nearest road or nearest urban centre (Lilongwe, Zomba or Blantyre) in kilometers.

Access to phone This is a dummy variable that equals one if there is a place in the village to make a phonecall, and zero otherwise.

Sex ratio This is the ratio of men to women at the cluster 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.

Size of cluster This is the total number of households in that particular household's marriage market cluster, defined using the method described in Section 3.2.

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 B: 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.

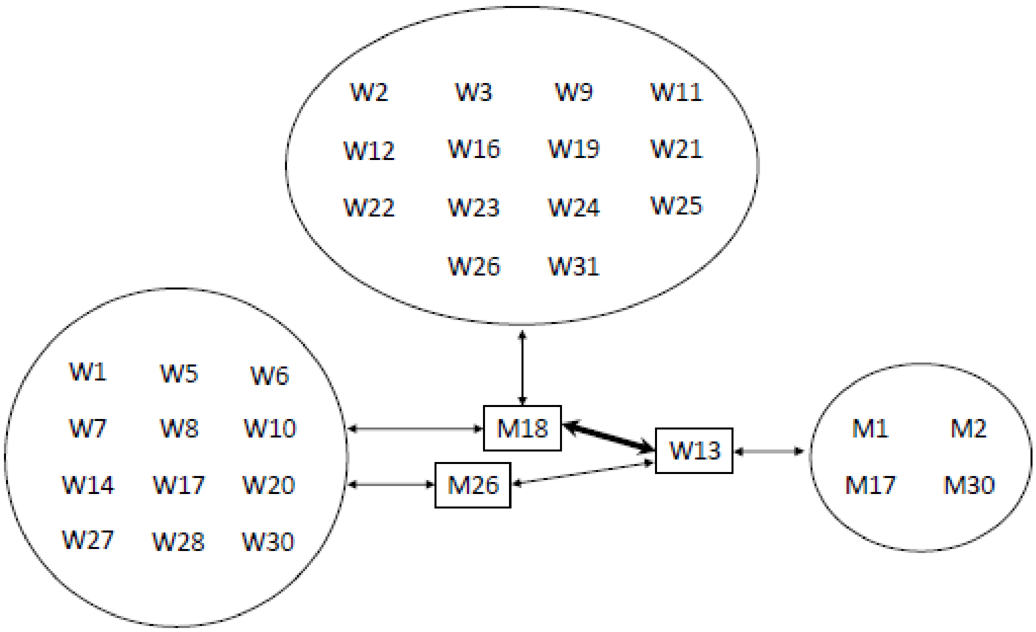

Figure 1: The instability network of a cluster. Arrows depict blocking pairs, M refers to men and W to women