Marriage and the City: Search Frictions and Sorting of Singles*

Pieter A. Gautier, VU University Amsterdam
Michael Svarer, University of Aarhus
Coen N. Teulings, University of Amsterdam, CPB

August 29, 2009

Abstract

This paper develops and tests a model where cities play an important role as marriage markets. The idea is simple. Cities are dense areas where singles can meet more potential partners than in rural areas. To enjoy those benefits, they are willing to pay a premium in terms of higher housing prices. Once married, the benefits from meeting more potential partners vanish and married couples move out of the city. Attractive singles benefit most from a dense market and are therefore more likely to move to the city. Those predictions are tested and confirmed with a unique Danish dataset.

1 Introduction

Why do people live in cities? The idea we put forward in this paper is that cities play an important role as marriage markets. Cities are dense areas where singles can meet more potential marriage partners than in rural areas. To enjoy these benefits, they are willing to pay a premium in terms of higher housing prices. Once married, the benefits from

*Acknowledgement: We thank the editor William Strange and two anonymous referees for their very useful coments and suggestions. We also thank Helena Skyt Nielsen and seminar participants at the Free University Amsterdam, Tinbergen Institute, University of Aarhus, University of Copenhagen, CEMFI, Copenhagen Business School, CREST, IZA, Singapore Management University, University of Helsinki, University of Utrecht and participants at the conference on Labour Market Models and Matched Employer-Employee Data in Honour of Dale T. Mortensen 2004 for very useful comments. Michael Svarer thanks the Danish National Research Foundation for support through its grant to CAM and Pieter Gautier thanks the NWO for support through a VIDI grant.

meeting more potential partners disappear and consequently, the countryside becomes more attractive. This generates a flow of married couples out of the city.

We find that the fraction of individuals living in one of the 5 largest cities in Denmark at age 18 is 22%, at the date of marriage this fraction has increased to 36% and after 5 years of continued marriage it bounces back to 23%.

In order to explain this pattern, we extend the marriage-market model of Burdett and Coles (1997) and distinguish between search markets that are more efficient (cities) and less efficient (rural areas). One obvious implication of the model is that singles are more likely to move from rural areas to cities while couples are more likely to make the reverse movement. This is in particular relevant for the most attractive types because they are most choosy and therefore benefit most from a high contact rate. Consequently, they are also the ones that are most willing to pay the higher house prices in the city. In a segmented equilibrium, all agents have a desire to sort in homogeneous segments and the joint existence of cities and rural areas offers them this opportunity. Our story can also be rephrased as one of clubs that sort attractive types and raise high entrance fees, see Jacquet and Tan (2007). Since the opportunity cost of being single are larger for the most attractive types, they are willing to pay those high fees and a separating equilibrium will result irrespective of the contact technology. Finally, Eeckhout (2006) discusses sorting on pay-off irrelevant characteristics.

We use canonical correlations to create attractiveness indices which are a linear combination of education, income, father's education and father's income. We find that (1) singles are more likely to move from the countryside to the city than couples, (2) couples have the largest probability to make the reverse movement and (3) attractive singles are more likely to move to the city than less attractive singles. We also test the sensitivity of our results to different definitions of attractiveness and cities. Moreover, we take subsamples of (i) individuals older than 25 to eliminate a potential college-effect and of (ii) individuals who never have kids to control for the possibility that children influence the moving decision. Our main results are robust to those exclusions, although our findings are less precise in the reduced sample for individuals without children. Finally, we are worried that our results are driven by life cycle motives. When people get older they enjoy clubs and bars less and walking through nature more. To address this issue, we consider individuals who divorced in the country. If their location choice was driven by life-cycle considerations we expected them to stay in the country while according to the marriage-market story we expect them to move back to the city. We give evidence that divorcees tend to move back to the city.

A number of papers are related to ours. Mincer (1978) argues that marriage reduces mobility because the cost are higher for families. He finds support for this pattern in US data.

Costa and Kahn (2000) argue that higher educated couples (power couples) are over-represented in cities. The idea is that the colocation problem (both have to live close to their job) is particularly severe for them. Their model therefore predicts that higher educated couples are more likely to move into the city and less likely to move out of the city. In terms of the latter prediction, our model predicts exactly the opposite. They use cross-section data from the U.S.. Recently, Compton and Pollak (2007) took another look at the issue. They argue that another explanation for the overrepresentation of power couples in the large cities is that all college educated individuals, married and unmarried, are attracted to the amenities and high returns to education of the large cities. As a result of this, the formation of power couples is more likely to occur in larger than smaller metropolitan areas. In their explanation education is key while in ours, the marriage market role of cities is the driving force. Based on PSID data, they analyze the dynamic patterns of migration, marriage, divorce, and education in relation to city size and find that power couples are not more likely to migrate to the largest cities in the U.S. than part-power couples or power singles. Instead, the location trends are better explained by the higher rate of power couple formation in larger metropolitan areas. With the Danish data we find that the marriage market role of cities is more important than the colocation of job opportunities. High skilled singles move to the city but once they are married they are more likely to move out of the city. Dahl and Sorenson (2008) give evidence that migration in Denmark is more driven by social factors (location of family and friends) than by financial incentives.

Stark (1988) and Smith and Thomas (1998) also look at the relation between marriage and mobility. Stark argues that the labor and the marriage market interact because the location where people search for jobs is often the same as where they search for a marriage partner. Smith and Thomas find for Malaysia that migration by males is mainly driven by labor market considerations while for females, fertility and family considerations are more important. We also find that females behave more like the marriage-market model than males.

Black et al. (2002) suggest that the reason a city like San Francisco hosts a disproportional high number of gays is due the high housing cost of living there. San Francisco is known as one of America's loveliest cities. Hence, due to the high demand for housing in San Francisco, housing prices are high. Gay couples face constraints that make having children more costly for them than for similar heterosexual couples. This frees resources for other "goods" such as housing in high-amenity locations. Although we do not explicitly consider the gay mating market, our model suggests an alternative explanation. Since the market for gays is relatively thin, they gain a lot by moving to a dense market like cities. In addition, any area that happens to have a large gay community will attract more gays because the matching rate depends not only on the contact rate but also on the share of potential mates and this is what pushes up house prices and creates nice amenities.

Edlund (2005) argues that young women outnumber young men in urban areas. The argument is that urban areas offer skilled workers better labor markets. Assuming that there are more skilled males than females, this alone would predict a surplus of males. However, the presence of males with high incomes may attract not only skilled females but also unskilled females, and thus a surplus of females in urban areas from the combination of better labor and marriage markets.

Finally, clustering and partner formation is studied intensively in biology. In many promiscuous species, females cluster around the top males. Secondary males may however also cluster around the top males for two reasons. According to the "hotshot" model of Beehler and Foster (1988), remaining isolated is less attractive since then they would meet even fewer females. Alternatively, Bradburry (1981) argues that males respond to female's desire to compare. In the presence of search frictions, clustering is necessary for comparison. Finally, Wagner (1997) gives evidence that socially monogamous species have similar incentives to cluster.[1]

The paper is organized as follows. First, in section 2 we present a simple marriage market model. In section 3 we discuss the data. Section 4 presents the main estimation

[1] Socially monogamous species are characterized by the fact that both partners engage in parental care. However, there is lot of evidence that both males and females still pursue extra-pair copulations (EPS). This EPS, give according to Wagner similar incentives to cluster as the motivations described above.

results. Section 5 carries out a number of robustness checks and section 6 concludes.

2 The model

The marriage market that we consider is in the spirit of Burdett and Coles (1997). Our economy is made up of two locations, the countryside and the city. All males and females have identical preferences but they differ in their attractiveness as a marriage partner. The attractiveness distribution and preferences are the same for both sexes. We assume that all singles are randomly distributed over the city and the country side at the age that they enter the marriage market.

We discuss the marriage decision problem from the female point of view, the male perspective is mutatis mutandis the same. Consider a single female who is looking for a marriage partner. Given the location choices of all the other singles, and the excess cost of living in the city, she chooses the location of residence that maximizes her expected discounted utility. For reasons of tractability, we analyze the model under two different extreme assumptions for the cost of mobility. Either this cost is zero for singles and finite for couples, or it is infinite for both. Both extremes allow us to focus on a particular aspect of the model, the sorting of individuals across locations (for zero mobility cost of singles) and the tightness of the correlation between male and female attractiveness within a location (for infinite mobility cost).

Each woman searches for a husband who she wants to marry *and* who wants to marry her. After having found a partner the couple must decide whether to stay at the current location or move. We treat the moving cost for couples to be a random variable which captures the idea that it depends on work location or the presence of children etc. The size of this moving cost is revealed after marriage. Divorce is ruled out, marriage is an absorbing state. We focus on the steady state where the yearly inflow of singles is equal to the yearly outflow of married couples and where the death rate is δ. For simplicity, we assume that both partners of a couple die simultaneously. To keep thinks simple, we set the rate of time preference at zero. In order to keep the model tractable we ignore labor market considerations. This is not because we believe that employment motivated migration is unimportant but rather because we want to focus here on the marriage market role of the city and there already exists a large literature on labor markets, agglomeration and the urban wage premium[2]. In our empirical analysis we do discuss how we can distinguish between both stories.

Let a be the attractiveness of a female and α the attractiveness of male, a and $\alpha \in [a^-, a^+]$, where we assume $a^+ > a^- > 0$. For simplicity we assume a to be constant over time. Coles and Francesconi (2008) allow a to decrease over time. Symmetry implies that the female problem is identical to the male problem. Let $l \in \{0, 1\}$ be the location of residence of an individual or a couple (0 = countryside, 1 = city) and let c be the excess cost of living in a city which we take as given but later in 2.4 we endogenize c for a special case of the model. Utility is non-transferable between males and females. It is convenient to discuss the model in reverse order, starting from the last stage of the life cycle, the location choice of a married couple, and then work back to the location choice of singles.

Lifetime utility of a married couple at the optimal location

A female who is married with a male who has attractiveness α and who lives in area l enjoys life time utility

$$u_m(\alpha, l) = \alpha - lc.$$

where c is the excess cost of living in a city (in stock terms). [3]

Location choice of a married couple. A couple chooses the location of residence that maximizes their utility. They trade off the cost of moving with the utility gain that a change of location yields. Denote the cost of moving by γ. We assume γ to be a strictly positive random variable. Couples learn the value of γ after marriage. A couple that marries in the city, $l = 1$, decides to move to the countryside, $l = 0$, if its moving cost are lower than the excess cost of urban life:

$$u_m(\alpha, 1) < u_m(\alpha, 0) - \gamma \Rightarrow c > \gamma, \tag{1}$$

A couple who marries in the countryside never moves to the city because the cost of living are higher there. The additive separability of the utility function in the attractiveness of the partner and the location specific cost makes that the preference of both husband

[2] See e.g. Rosenthal and Strange (2004), Gautier and Teulings (2004, 2009), Combes, Duranton and Gobillon (2008) and Gemici (2008).

[3] Utility depends only on the characteristics of one's partner. Specifications where the own type matters in an additive way do not change the results, i.e. a female with attrictiveness a^f married to a male a^m receiving utility, $u = a^m + h(a^f)$, with $h' > 0$. In the words of Burdett and Coles (1999): "narcissm is not necessarily ruled out".

and wife for living at either location is the same. Hence, we can ignore any problem of intra-household bargaining on location choice. In case of infinite mobility cost, married couples never move.

Lifetime utility of a single and a married female. Define the expected utility for a female who is married in location l to a male with attractiveness α as[4]:

$$\begin{aligned} \mathrm{E}_l u_m(\alpha, l) &\equiv \mathrm{E}_l \left[\max\left[u_m(\alpha, 1), u_m(\alpha, 0) - \gamma\right]\right] = \alpha - l(c - C), \\ C &\equiv \mathrm{E}\max\left[c - \gamma, 0\right]. \end{aligned}$$

C is the expected cost saving of moving out of the city. C is equal to either zero if the moving cost exceeds the excess cost of living in the city or it is equal to the difference between the cost of living in the city and the moving cost if the latter are less than the former.

Let $f_l(\alpha)$ be the mass of males of attractiveness α who search for a partner in location l, and $m_l(a)$ the marriage set of a female with attractiveness a. This set consists of all males with attractiveness $\{\alpha\}$ with whom she is willing to marry and who are willing to marry her. Hence:

$$\alpha \in m_l(a) \Leftrightarrow a \in m_l(\alpha).$$

The lifetime utility for a single female who searches for a partner in location l is determined by the following Bellman equation:

$$\begin{aligned} \delta u_s(a, l) &= \lambda_l \int_{m_l(a)} \left[\mathrm{E}_l u_m(\alpha, l) - u_s(a, l)\right] f_l(\alpha)\, d\alpha - \delta l c \\ &= \lambda_l \int_{m_l(a)} \left[\alpha - l(c - C) - u_s(a, l)\right] f_l(\alpha)\, d\alpha - \delta l c. \end{aligned} \tag{2}$$

The death rate, $\delta > 0$, plays the same role as the discount factor in standard search models. λ_l is the arrival rate of marriage candidates per unit of the stock of searching candidates. We assume that the arrival rate is higher in the city than in the country, $0 < \lambda_0 < \lambda_1$, either because of a higher density or because the market is bigger. The first right hand term of the first line is the welfare gain of marriage and moving to the optimal location. The second term, $\delta l c$, is the excess cost of living in the city that a single pays during the period she searches for a partner. Note that the assumption $a^- > 0$ implies

[4] Where we can think of $\alpha \equiv \alpha'/\delta$ with α' being the flow value of marriage with a type α.

that a woman who lives in the countryside prefers being married to the least attractive male to remaining single for ever, since remaining single yields $u_s(a,0) = 0$, while being married to the least attractive man yields $\mathrm{E}[u_m(\alpha,0)] = a^- > 0$.

Marriage sets. Marriage requires mutual agreement. Hence, for both partners, the lifetime utility of being married must be weakly greater than the lifetime utility of being single.

$$\begin{aligned} C1 &: \mathrm{E}_l u_m(\alpha,l) > u_s(a,l) \wedge \\ C2 &: \mathrm{E}_l u_m(a,l) > u_s(\alpha,l). \end{aligned} \tag{3}$$

Condition $C1$ states that a female with attractiveness a must be willing to marry a male of attractiveness α, $C2$ states that a male with attractiveness α must be willing to marry a female of attractiveness a. By symmetry, the marriage sets of a man and a woman with the same attractiveness are the same.

Equality of the in and outflow of singles for each a and l. We assume that the marriage market in each region is in steady state. Hence, the number of new singles entering the market equals the number of singles getting married. Let $g_l(a)$ be the mass of singles of attractiveness a entering location l to search for a partner. In steady state the following holds:[5]

$$g_l(a) = f_l(a) \times \left(\lambda_l \int_{m_l(a)} f_l(\alpha)\, d\alpha + \delta\right). \tag{4}$$

The left hand side is the inflow of new singles of attractiveness a, and the right hand side is the outflow rate which equals the number of contacts between females with attractiveness a with males in her matching set, $\int_{m_l(a)} f_l(\alpha)\, d\alpha$ and a fraction δ dies. By symmetry, a similar equation holds for the inflow of males of attractiveness α, $g_l(\alpha)$ and the mass of males and females with the same attractiveness, $f_l(a)$ is identical. In appendix A.1 we derive a closed form expression for $f_l(a)$

Location choice of a single. For the case of costless mobility of singles, a female living in the countryside, $l = 0$, decides to move to the city, $l = 1$, if

$$u_s(a,1) > u_s(a,0). \tag{5}$$

[5] Hence, we deviate from a common but unpleasant simplification in the literature which is the cloning assumption: each person who gets married is immediately replaced by another person of the same attractiveness, see e.g. Bloch and Ryder (2000). That assumption fixes the distribution of attractiveness over the *stock* instead of over the *inflow*.

and vice versa. Let $h(a)$ be the mass of singles of attractiveness a in both locations before location choice and let $I(a)$ be an indicator function that takes the value one if attractiveness level a prefers location 1, that is, if condition (5) is satisfied and zero otherwise. Then

$$\begin{aligned} g_1(a) &= I(a)\,h(a), \\ g_0(a) &= [1 - I(a)]\,h(a). \end{aligned} \tag{6}$$

When the cost of mobility is infinite (see section 2.3), we assume the new singles to be equally distributed among both locations:

$$g_l(a) = \frac{1}{2} h(a). \tag{7}$$

Definition 1. *An equilibrium is a collection of marriage sets* $m_l(a)$and *densities* $f_l(a)$ *that satisfies equations (3) and (4), and either equation (6) for the case with zero mobility cost for singles, or equation (7) for the case with infinite mobility cost.*

Below, we first characterize the equilibrium for a single location. Next, in 2.2. we turn our attention to the sorting of singles and couples into cities and rural areas in the case of zero mobility cost for singles. Then, we compare the correlation between male and female attractiveness between both locations for the case of infinite mobility cost in section 2.3. Finally, we explain in section 2.4 why cities are more expensive to live in, have a higher contact rate and have a larger fraction of singles.

2.1 The marriage market equilibrium

The shape of the marriage sets is determined by a number of simple observations. If a female with attractiveness a is willing to marry a male with attractiveness α^*, then she is also willing to marry all males who are more attractive: $\alpha > \alpha^*$ because $u_m(\alpha, l)$ is strictly increasing in α. Hence, the marriage set of a woman with attractiveness a in location l is convex, the lower bound $a_l^-(a)$ being the least attractive man to whom she is willing to marry, the upper bound $a_l^+(a)$ being the most attractive man who is willing to marry her. In other words, the lower bound is the attractiveness α for which condition $C1$ is just violated (i.e. holds by equality), the upper bound is the highest rank α for which condition $C2$ is not violated (i.e. holds by equality). Hence, the marriage set of a woman with attractiveness a is defined as $\{\alpha\} \in \langle a_l^-(a), a_l^+(a)]$.

Consider the most attractive female, $a = a^+$. By the previous argument, all males are willing to marry her, so the upper bound of her marriage set is $\alpha = a^+$. The lower bound of her marriage set, denoted a_l^1, is determined by the male type that gives her the same amount of marriage utility as the value of remaining single. Evaluating (2) at α_l^1 gives an expression for $u_s(\alpha_l^1, l)$. Then,

$$\begin{aligned} a_l^1 - l(c - C) &= u_s\left(a_l^1\right) = \frac{\lambda_l}{\delta}\int_{a_l^1}^{1}\left(\alpha - a_l^1\right) f_l(\alpha)\, d\alpha - lc \\ a_l^1 &= \frac{\lambda_l}{\delta}\int_{a_l^1}^{1}\left(\alpha - a_l^1\right) f_l(\alpha)\, d\alpha - lC. \end{aligned} \tag{8}$$

It is easily verified that all single women with attractiveness $\{a\} \in \langle a_l^1, 1]$ set the lower bound of their marriage set at the same value as the most attractive woman because they solve exactly the same problem. Hence, all these single women have the same utility, $u_s(a, l) = u_s(a^+, l), \forall a \in \langle a_l^1, a^+]$. By symmetry, the same applies to all males. The females and males with attractiveness $\{a\} \in \langle a_l^1, 1]$ form a closed segment and marry with each other, but they do not marry with anybody else. A woman with attractiveness a_l^1 can therefore not marry with a more attractive male. Her own attractiveness is the upper bound of her marriage set. The whole logic that applies to the first segment that contains the most attractive candidates therefore carries over to the second segment. The lower bound of the next segments i, a_l^i, can be calculated in a similar way:

$$a_l^i = \frac{\lambda_l}{\delta}\int_{a_l^i}^{a_l^{i-1}}\left(\alpha - a_l^i\right) f_l(\alpha)\, d\alpha - lC. \tag{9}$$

The whole market falls apart in a number of consecutive, non overlapping segments. Men and women marry within and never outside their segment. In A.2 we characterize the equilibrium and give a condition for uniqueness. Figure 1 shows the segments in the attractiveness space of males and females. The segments are given by the shaded areas. By equation (19), a female's attractiveness when being single is directly related to the utility of the least attractive male in her segment, because all females in this segment are indifferent between marrying this man and remaining single. Hence, this utility can be derived immediately from Figure 1. This segmented class structure has been established in a number of papers.[6] In a "Walrasian" marriage market without search cost, the utility

[6] See e.g. Burdett and Coles (1997), Bloch and Ryder (2000), Eeckhout (1999), and Smith (1997).

of a single woman equals her attractiveness, because each level of attractiveness forms a separate segment and marriage sets are reduced to singletons. Consequently, all matches are on the diagonal (Gale and Shapley, 1962). Hence, for the countryside (for the city the argument is only trivially different) the vertical distance between the diagonal and the actual utility $u_s(a,0) = a_0^i, a_0^i < a < a_0^{i-1}$ can be interpreted as a measure of the cost of search frictions for a woman of attractiveness level a. Only for the least attractive single woman in each segment her utility is equal to what it would be in a "Walrasian" market. For her, the cost of waiting for a suitable marriage partner is exactly offset by the chance of finding a better partner than she would have been able to find in a "Walrasian" market. A slight change in the segmentation would therefore make her worse off, since she would no longer be the least attractive woman in her segment, and hence she would get a lower pay off than in the "Walrasian equilibrium ". Hence, there is no unambiguous Pareto ranking for equilibria with a different number of segments. The shaded area between the diagonal and the actual utility is a measure of the total cost of search frictions. The larger therefore the number of segments, the smaller the total losses due to search for the most attractive type are.

2.2 The best location to search for a partner

Having characterized the marriage market equilibrium in each location, we can now proceed with the analysis of the location choice at the moment that a single starts looking for a partner. The case with infinite mobility is simple: there is no mobility, everybody starts searching for a partner at the location where she is born. With finite mobility cost for couples and zero mobility cost for singles, the latter move to the location where their expected utility is highest. In that case, there are multiple equilibria.[7] The location choice determines the density of new singles that enter the market in a particular location, $g_l(a)$, which in turn determines the number of people with attractiveness a in the stock of people who are looking for a partner, $f_l(a)$. An extreme example clarifies the problem. For the sake of argument, assume $c = 0$ at this point. Suppose that all singles decide to look for a partner in the city, then $g_0(a) = f_0(a) = 0$ for all a, and hence, $u_s(a,0) = 0$, which rationalizes the choice of singles to move to the city in the first place. Clearly,

[7] Formally, this can be seen by comparing the flow equilibrium condition (16) in the appendix to equation (6) for $g_l(a)$.

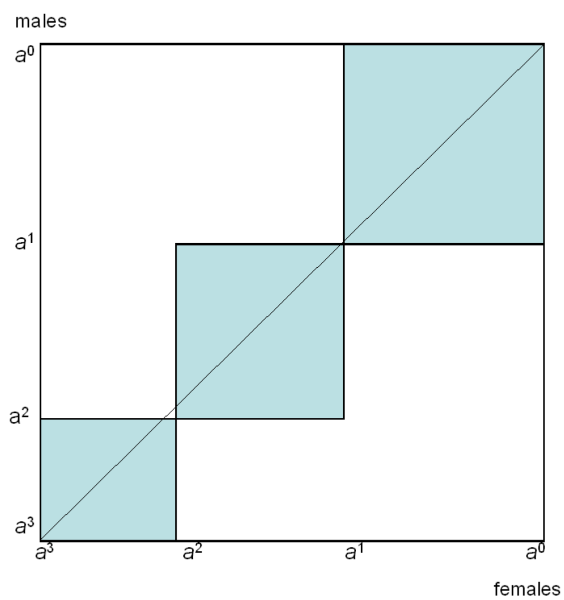

Figure 1: Marriage market segments

when we start with the reverse presumption that all singles decide to look for a partner at the countryside, then $g_1(a) = f_1(a) = 0$ for all a, and hence, $u_s(a, 1) = 0$, which also rationalizes the presumption. This explains the existence of small areas like Beverly Hills which contains a large fraction of attractive singles. If we want to obtain a unique equilibrium, we have to impose further restrictions. For this purpose, we introduce the notion of hierarchical efficiency.

Definition 2. *An hierarchically efficient equilibrium is an equilibrium that satisfies Definition 1 and where on top of that the lifetime utility of a single with attractiveness* a^* *cannot be improved without making a more attractive single* $a \in \langle a^*, a^+]$ *worse off.*

By the class structure of the equilibrium on the marriage market, where a single in the i-th segment of the market has to take the behavior of singles in higher segments as given, the concept of hierarchical efficiency is related to Aumann's (1953) concept of a *strong equilibrium* (there exists no profitable deviation by a coalition of players). In an hierarchically efficient equilibrium, the utility of each individual is maximized subject to the constraint that the utility of more attractive people is already maximized. Since in any equilibrium, a coalition of less attractive people cannot influence the strategy of more attractive people, an hierarchical efficient equilibrium is a strong equilibrium.

Proposition 2. *Consider the case of finite mobility cost. For intermediate values of* C*, there exists a critical level of attractiveness* a^* *such that in a strong equilibrium all singles with* $a \leq a^*$ *move to the countryside,* $l = 0$ *and all others to the city,* $l = 1$*. For lower values of* C*, everybody moves to the city. For higher values of* C *everybody moves to the countryside.*

Proof. See appendix A.3.

Proposition 2 shows that the most attractive single women prefer the city and the least attractive women prefer the countryside. We label this outcome the *elite city ordering.* Singles face a trade off between the efficiency of the marriage market in the city and the cheap cost of living in the countryside. The elite prefers the efficiency of the city marriage market above the cheap cost of living in the country because they gain more by a higher contact rate. Their greater attractiveness allows them to marry with more attractive partners since their own attractiveness is basically their endowment in the marriage market. Hence, their opportunity cost of remaining single are the highest and therefore they have the greatest incentive to move to the city. We have not modelled the housing market at this stage. However, one can expect that with a positive supply of housing both in the city and at the countryside, house prices will adjust such that at least some singles prefer the city above the countryside, for otherwise nobody wants to live in the city (remember that couples migrate only from the city to the countryside), see section 2.4.

Tan and Jacquet (2008) study a related problem where market places fulfill the same role as locations in our model. They allow for free entry of market places and show that all equilibria feature perfect segmentation (the first contact results in a match) when the contact technology exhibits constant returns to scale. With a quadratic contact technology they show that (like here) a Burdett Coles class structure arises. The intuition is that under CRS, less attractive types impose congestion externalities on more attractive types and this gives the latter more incentives to form new market places for themselves. Under a quadratic technology, there are no congestion externalities.

2.3 Location differences in the tightness of matching

Next, we consider whether the segmentation of the marriage market is tighter in the city. With finite mobility cost the analysis is very messy as we argue below and therefore we focus on infinite mobility cost in this section and assume that the distribution of attractiveness of the inflow of new singles is equal in both regions. At what location is the correlation between the male and female attractiveness of a couple highest?

Proposition 3. *Consider the case of infinite mobility cost. Conditional on a common upper bound a^{i-1} of segment i at both locations, the lower bound of the segment is lower in the country than in the city $a_0^i < a_1^i < 1$.*

Proof. This follows from the fact that the lower bound of a segment is increasing in λ, see (9).■

Since the contact rate is higher in the city, singles respond by becoming more choosy and this implies a higher correlation between male and female attractiveness since matches are closer to the main diagonal of Figure 1.

The relation between segment size and regions is more complicated with finite random mobility cost. In that case, part of the singles of a particular level of attractiveness will move and part will stay at the location where they are born (those with high mobility cost). It then depends on the exact segmentation of the market whether singles of a particular level of attractiveness move to the city or to the countryside. For example, the best segment, $i = 1$, will be tighter in the city than at the countryside, but a female with attractiveness a in between the lower bounds of the upper segment in the city and at the countryside falls in the second segment in the city while she belongs to the first segment in the country. If the lower bound of this second segment in the city is below the lower bound of the first segment in the countryside, ($a_1^1 < a_0^1 < a_1^2$), she remains in the country. Hence, there is myriad of potential outcomes.

However, with finite random mobility cost, one can expect by and large that for the upper end of the attractiveness distribution there is net immigration to the city and matching is tighter there, while for the lower end of the distribution, the reverse holds. Immigration and tighter matching are two sides of the same coin, as tighter matching implies that the most attractive singles in a segment get higher utility in that location, and hence singles will move there. As we will show in the next subsection, the higher cost of living in the city can be driven by the net aggregate migration of singles to the city.

2.4 The nature of the excess cost of living in the city

This section proposes an explanation for why living in the city is more expensive than in the countryside. For reasons of tractability, we consider a simplified version of our model where all agents have the same attractiveness a and where the stock of housing (and hence the population) at both locations is equal. We normalize the total population size to one and start our analysis with equal contact rates in both locations, $\lambda_1 = \lambda_0 = \lambda$. Without loss of generality we denote the location with (weakly) more singles as the city, $l = 1$.

Contrary to the previous analysis the excess cost of living in the city c is now endogenous. Since we will show that the cost of living in the city are equal to or above the cost of living in the country, $c \geq 0$, couples who married in the country never move to the city.

Let $\Gamma(\gamma)$ be the distribution function of moving cost, γ. Couples move to the countryside if $\gamma < c$. So a fraction $\Gamma \equiv \Gamma(c)$ of the couples moves from the city to the countryside. Let F_i be the stock of singles in location i. The stock of married living in the country, M_0, follows from the following steady state flow condition:

$$\lambda F_1^2 \Gamma + \lambda F_0^2 = \delta M_0.$$

The left hand side is the inflow of married couples into the country; where the first term is the inflow of the newly married couples with sufficiently low moving cost moving in from the city, and the second term represents the number of marriages in the country. The right hand side is the outflow of married couples (into death). The stock of married in the city, M_1, follows from a similar equation:

$$\lambda F_1^2 (1 - \Gamma) = \delta M_1.$$

The stock of married couples at the countryside, M_0, plus the stock of singles, F_0, must be equal to half of the total population. The same applies to the city. Hence:

$$\frac{\lambda}{\delta} F_1^2 \Gamma + \frac{\lambda}{\delta} F_0^2 + F_0 = \frac{1}{2}, \tag{10}$$

$$\frac{\lambda}{\delta} F_1^2 (1 - \Gamma) + F_1 = \frac{1}{2}. \tag{11}$$

The lifetime utility of a single living in respectively the country and the city is given by:

$$\begin{aligned} u_{s0} &= \frac{\lambda}{\delta} F_0 (a - u_{s0}), \\ u_{s1} &= \frac{\lambda}{\delta} F_1 (a - u_{s1}) - (c - C). \end{aligned}$$

Since singles are free to choose their location, a no-arbitrage condition holds and both utilities must be equal in equilibrium, $u_{s0} = u_{s1}$, implying that:

$$c - C = \frac{\lambda (F_1 - F_0)}{\delta + \lambda F_0} a, \tag{12}$$

Suppose γ is distributed uniformly on the interval $[0, 1/\eta]$. This implies:

$$\begin{aligned} \Gamma(\gamma) &= \eta\gamma, \Gamma = \eta c, \\ C &\equiv \mathrm{E} \max[c - \gamma, 0] = \Pr(\gamma < c) \cdot E(\gamma | \gamma < c) = \frac{1}{2} \eta c^2. \end{aligned}$$

Using these conditions, equilibrium can be characterized by an equation that only depends on c.[8] There exist multiple equilibria. The first is a trivial equilibrium where both

[8]

$$0 = \frac{\left(2\lambda F_1 + \delta - \delta\sqrt{1 - 4\left(\frac{\lambda}{\delta}\right)^2 F_1^2 \eta c + 2\frac{\lambda}{\delta}}\right)}{\delta + \delta\sqrt{1 - 4\left(\frac{\lambda}{\delta}\right)^2 F_1^2 \eta c + 2\left(\frac{\lambda}{\delta}\right)}} a - c + \frac{1}{2}\eta c^2, \tag{13}$$

$$\text{where } F_1 = \frac{-\delta + \delta\sqrt{1 + 2\left(\frac{\lambda}{\delta}\right)(1 - \eta c)}}{2\lambda (1 - \eta c)} \tag{14}$$

where (13) follows from solving (10) for F_0 and substituting the result in (12) while (14) follows from solving (11) for F_1.

locations have the same population composition and where $c = 0$. Since moving cost are positive, married couples have no incentives to move out of the location where they marry. For sufficiently low moving cost, this equilibrium is unstable because if the initial conditions are such that the city has slightly more singles, it will break down and the city attracts more singles. Then, the city specializes in marriage formation and consequently it attracts relatively many singles who push up house prices and increase the cost of living. This gives newly formed couples with sufficiently low moving cost incentives to leave the city in order to save on the high cost of living. Below, we illustrate the equilibrium for $\{\eta = 1, a = 1, \lambda = 1, \delta = 0.2\}$.

Besides the equilibrium where $c = 0$, there is one other equilibrium in this example where $c = 0.52$. Location 1 has more singles than location 0, $F_1 = 0.29, F_0 = 0.16$ and location 0 has more married couples: $M_1 = 0.21$ and $M_0 = 0.34$. We have not found parameters combinations that generate more than 2 feasible equilibria. The higher cost of living in the city are due to the rents that home owners appropriate. Suppose that we extend this model with an upward sloping housing supply curve. The idea is that construction becomes more and more expensive as the number of houses in an area increases. Since rents are higher in the city, equilibrium housing supply will be higher in the city as well. Consequently, the population density in the city will be larger than in the country, raising the contact rate in the city above the contact rate at the countryside, $\lambda_1 > \lambda_0$, and making the city even more attractive for singles. This explains why the excess cost of living in the city go hand in hand with a higher population density. Finally, the supply of single amenities (bars, clubs etc.) will also adjust to the composition of a region. In our model, the emergence of cities is the result of the desire of singles to cluster together in order to find an attractive marriage partner. Adding heterogeneity to this model makes that cities specialize in their role as efficient marriage markets in particular for attractive singles, as discussed in Section 2.2.

3 Empirical strategy and data

In the remainder of the paper we provide a simple exploratory empirical analysis of the main predictions of the theoretical model presented above. Our goal is to test whether the predictions are consistent with a data set consisting of a randomly drawn subsample of Danes born between 1955 and 1965. Specifically, we look at the moving patterns between the countryside and cities and we investigate whether singles and in particular the most attractive ones are more likely to move to the city. We consider some alternative explanations in Section 5. Below, we give a description of the data set used and how we construct a measure of attractiveness. In the next section we present the empirical results which are based on linear probability models of the mobility pattern.[9]

The data that we use to test the main implications of the model come from IDA (Integrated Database for Labor Market Research) created by Statistics Denmark. The information comes from various administrative registers that are merged at Statistics Denmark. The IDA sample used here contains (among other things) information on marriage market conditions for a randomly drawn sub-sample of individuals born between January 1, 1955 and January 1, 1965.[10] The individuals are followed from 1980 to 1995. The data set enables us to identify individual transitions between different states on the marriage market on an annual basis. In addition, we have information about current geographical location. This implies that we also observe an individual's mobility pattern within Denmark on an annual basis.[11] If the individual enters a relationship we also observe the personal characteristics of the partner. There are 23672 individuals in our sample. We use the following variables:

Education. We use years of completed education to describe educational attainment. Since most of the sample is acquiring education in the sample period we will use education in 1995 (when the youngest person in the sample was 30) as the indicator for level of education (to avoid problems with unfinished education). The basic level of education in Denmark is 9 years of education. Individuals graduating from high school have 12 years of education, the next education class is 14 years of education, to complete a medium degree requires 16 years and finally it takes 18 years to graduate from university.

Income. We use (log) gross income. The income figures are all in terms of 1980 prices and includes salary, capital gains, and income transfers. The consumer price index is used as a deflator.

[9] In earlier versions of the paper we estimated non-linear probability models, which gave very similar results. Since one of our main variables is the interaction term between being single and level of attractiveness we find it more informative to use the linear probability model which gives a cleaner interpretation of the interaction term.

[10] Immigrants are included but cannot be separately identified.

[11] We consider Denmark to be a single market. Given the small fraction of immigrants this is not too restrictive (the foreign born population only accounted for 5% in 2000 while for the OECD as a whole this is 10%, see OECD 2008) .

City- rural definition. We divide Denmark into cities and rural areas. The five largest Danish cities are Copenhagen (incl. Frederiksberg), Aarhus, Odense, Aalborg, and Esbjerg.

The most dense area in Denmark is the Copenhagen metropolitan area. 12.7 % of the population lived there in 1995. The other cities host 15% of the population in 1995. The five cities are distributed across the country as shown in Figure 2. We therefore conjecture that the relevant city definition is to include the largest cities in each region of Denmark. The dependent variables in the analysis are indicator variables that take the value 1 if an individual that lives in the city moves to the countryside or, for the other regression, if an individual that lives in the countryside moves to the city in a given year. For individuals who stay either in the city or in the countryside two consecutive years the relevant dependent variables take the value 0. Note that a city according to our definition is a much smaller unit than a CMSA in the Census data so one should be careful when comparing our results to those of Costa and Kahn (2000) and Compton and Pollak (2007). Therefore, we repeated our analysis with a different definition of dense and non-dense areas based on the population density. This changes the city definition somewhat. Some suburbs of Copenhagen are more densely populated than the large cities. It turns out however, that our main results are robust to changes in the city definition. Finally, we do our analysis only treating Copenhagen as a city and find qualitatively similar results.

Figure 2: Map of Denmark

Marriage. Individuals can occupy one of three states in the marriage market: single, cohabiting, or married. In this paper we merge cohabitation and marriage into one group and refer to this group as married. Cohabitation as either a prelude to or a substitute of marriage is very common in Denmark (see e.g. Svarer, 2004). There are some qualifications to this definition of marriage. Some of the couples - presumably a small minority - that are registered as cohabiting are simply sharing a housing unit, and do not live together as a married couple.

Personal characteristics. In addition to the information presented above we also have detailed information about the number of kids in the household. We know the labor market status of the individuals, their age and their income. In addition, we have information on the income and education of the father of each individual in the sample.

Table 1 presents descriptive statistics for the main variables. The period covers 1980-1995. We report means for 1987 for the main explanatory variables. These are representative for the whole period.

Table 1: Descriptive statistics

	City		Rural	
	Women	Men	Women	Men
Number of observations, 1987-numbers	3612	3815	7920	8325
Single, 1987-numbers (%)	46.3	55.4	26.0	44.0
Children, 1987-numbers (%)	35.4	19.3	56.2	33.3
Age, 1987-numbers (in years)	26.4	26.7	27.0	26.9
Years of education	12.14	12.36	11.89	12.01
Gross income (conditional on working, in 1,000 dkk)	194.1	256.9	183.7	253.8
Father's years of education	10.56	10.46	10.24	10.06
Father's gross income (in 1,000 dkk)	150.1	109.0	124.8	104.6
Father has missing income (%)	35.7	34.8	35.4	31.7
Other descriptives				
1995 House prices per m^2 (in 1,000 dkk)	6.3 DKK		4.3 DKK	
Average annual rural-to-city mobility rate (%) (all years)			2.9	
Average annual city-to-rural mobility rate (%) (all years)	6.4			

Note: Numbers represent percentages - unless stated otherwise.

Table 1 shows that a higher fraction of the population is single in the city although the age difference between the city and rural areas was quite small in 1987. That is, the higher fraction of singles in the city is not a result of a lower mean age. More people have children in the rural areas. People tend to be more educated and have higher incomes in the city. In addition, housing prices are also higher in the city. Table 1 also reveals that men marry later because the cohort contains relatively many single men. The average annual mobility rate for the sample period is presented in the last two rows. On average 2.9% of the individuals move from the rural areas to the city per year. The reverse move from a city to a rural area happens more than twice as often.

In the next section we construct an attractiveness measure which we use in section 4 to test the main implications of the model.

3.1 Constructing a measure of attractiveness

The model presented in Section 2 suggests that more attractive singles are more likely to move to the city. Individual attractiveness presumably depends on a whole range of characteristics like weight, height, age, intelligence, humor, physical appearance, income etc. Data limitations restrict us from using a complete set of personal attributes. Regrettably, pictures of the individuals in the sample are not available, so that we cannot rank individual according to their looks, as in e.g. Hamermesh and Biddle (1994). We therefore follow Wong (2003) and Anderberg (2004) and use income and education as attractiveness components. In addition, we also exploit information on father's level of education and income.[12] Below, we explain how we determine their relative importance. In a frictionless world, the most attractive females marry the most attractive males, resulting in a perfect correlation between male and female attractiveness. In a world with frictions this correlation will not be equal to one but it will be positive. Here, we conjecture that attractiveness for both males (A^M) and females (A^F) is a linear function of the four factors described above and the dummies for missing income and education for fathers.[13]

$$\begin{aligned} A^M &= edu * \alpha_1 + \ln(inc) * \alpha_2 + f\_edu * \alpha_3 + \ln(f\_inc) * \alpha_4 \\ &\quad + f\_miss\_inc * \alpha_5 \\ A^F &= edu * \beta_1 + \ln(inc) * \beta_2 + f\_edu * \beta_3 + \ln(f\_inc) * \beta_4 \\ &\quad + f\_miss\_inc * \beta_5. \end{aligned}$$

We estimate the relative importance of those factors (the $\alpha's$ and the $\beta's$) by canonical correlation, as was already suggested by Becker (1973). Canonical correlations (see e.g. Johnson and Wichern (1998)) construct indices of A^M and A^F such that the correlation between each of them is maximized subject to the indices being orthogonal to each other. In the model, we assumed that the two sets of variables are related to each other only through a single index. In Table 2 we present the results from the canonical correlation analysis. We standardized all variables, i.e., they are all transformed to have mean 0 and variance equal to 1 to ensure that the weights are consistently estimated.

All estimated coefficients are significantly different from zero. The standardized coefficients show that the attractiveness level is mainly determined by education. The bivariate

[12] In labor economics similar issues arise, see e.g. Heckman and Scheinkman (1987). Worker skills are often measured by education but as Bacolod, Blum and Strange (2009) argue, this is restrictive. They use many other psychological traits and measures of intelligence and take an hedonic approach using information on job requirements to identify skills from skill requirements. Gautier and Teulings (2009) also take an hedonic approach and use all worker characteristics that are available in the CPS to construct a single dimensional skill index using wages to obtain weights for the various factors.

[13] We did not include age because it is likely that preferences are based on age differences between own's and partner's type rather than the absolute value of age. We also considered using occupation information. It was however not obvious how to rank different occuptations so we choose not to include it in the analysis.

correlation between the partners educational level is 0.38. The additional terms in the attractiveness measures do not add much to the total correlation. On the other hand they all have a significant impact on the measure and we choose to include them in the subsequent analysis. We also investigate a version where only education is included as a measure of attractiveness[14]. The first canonical root is 0.42 and although the second is significantly positive it is much smaller. Hence, the first canonical correlation captures most of the correlation between the two sets of variables and we can use a single index.

Table 2: Results from canonical correlations

	Canonical coefficients	t-value
α_1 : Man's education	0.86	44.55
α_2 : Man's father's education	0.23	10.31
α_3 : Man's income	0.13	6.79
α_4 : Man's father's income	0.35	2.70
α_5 : Man's father has missing income	0.32	23.50
β_1 : Woman's education	0.88	45.87
β_2 : Woman's father's education	0.24	11.23
β_3 : Woman's income	0.07	3.57
β_4 : Woman's father's income	0.32	2.80
β_5 : Woman's father has missing income	0.31	2.80
1. canonical correlation between A^M and A^F	0.42	
2. canonical correlation between A^M and A^F	0.14	
Bivariate correlations		
Education	0.38	
Income	0.07	
Father's education	0.14	
Father's income	0.15	
# couples	14018	

Note: All weights are significantly different from 0 at the 5% level.

Based on the estimated weights we construct an attractiveness number for each individual by adding up the weighted values of their characteristics. The summary statistics

[14]We also investigated a version of the model where we use average income in the last three years to purge for potential temporary income fluctuations. The results presented in Table 2 and in subsequent tables where not affected by this other income measure.

for the attractiveness index (singles only) for both cities and rural areas are presented in Table 3.[15]

Table 3: Summary statistics for the standardized attractiveness measure

	Singles	
	Mean	Std. Dev.
Rural		
Male attractiveness	-0.087	0.883
Female attractiveness	-0.085	0.902
City		
Male attractiveness	0.397	1.186
Female attractiveness	0.376	1.166

Table 3 shows that on average singles living in the city are more attractive than singles living in the countryside. Since, a crucial determinant of attractiveness in the current analysis is the level of education this might simply reflect that returns to education are larger in the city. In the next section we try to investigate whether other motivations for living in the city dominate the marriage market effect.

4 Results

In this section we investigate whether the mobility patterns observed in the data are consistent with the main predictions of section 2. First, singles have a relatively larger probability to move to the city and a relatively smaller probability to move to the countryside to explore the higher contact rate. Second, in particular the attractive singles have a larger probability of moving to and staying in the city since the opportunity cost of remaining single are largest for them.

In Table 4, we present the results for the movements rural to city and city to rural for both men and women. Table 4 shows that for both men and women the mobility pattern is consistent with the model predictions. That is, single people are more likely to move to the city and less likely to leave the city compared to married individuals. In addition, the coefficients of the interaction terms reveal that among singles the most attractive are most

[15] The attractiveness measure is constructed based on 1995 observations. The results throughout the paper are however unaffected if we use each year's income to construct the measure. Moreover, if we estimate the weights for city and rural areas separately, we get very similar results.

likely to move to the city and are also less likely to leave the city although for the city-to-rural movement, the sum of the coefficients (single, attractiveness and their product) is not significantly different from zero. Another interesting finding is that the presence of children is associated with a lower mobility propensity in all dimensions, although the marginal effect is, as expected, much larger in the mobility equation from the city to rural areas.

Table 4: linear probability model: probability of mobility[16]

	Women				Men			
	Rural to city		City to rural		Rural to city		City to rural	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Single	0.026*	0.002	-0.014*	0.003	0.007*	0.002	-0.015*	0.004
Attractiveness	0.005*	0.001	0.011*	0.002	0.009*	0.001	0.012*	0.002
Single*attract.	0.016*	0.001	-0.014*	0.003	0.013*	0.001	-0.007*	0.003
Children	-0.022*	0.001	-0.009*	0.004	-0.019*	0.001	-0.006	0.004
# Observations	84736		29714		76130		27065	

Note: *denotes significantly different from 0 at the 5% level.

An alternative story is that expected income is higher in the city (in particular for attractive types). However, higher income does not explain why singles and attractive singles in particular are more likely to move to the city than couples. Another competitor to the marriage market hypothesis of why attractive people locate in the city is that they can benefit from the higher returns to education. But this cannot explain why they move out of the city as we find. The finding that more attractive people are more mobile is consistent with other studies on individual mobility (e.g., Greenwood, 1997 and Compton and Pollak, 2007).

In Table 4, we also see that the probability of moving from the city to the countryside is in particular lower for the most attractive singles. That is, among the married the most attractive are more likely to leave the city. We believe this finding highlights the importance of considering the marriage market component in the investigation of mobility patterns between cities and rural areas. Our results are consistent with recent literature that looks at mobility patterns of individuals based on their marriage market status. Both Edlund (2005) and Compton & Pollak (2007) find that singles are more likely to locate in the cities and that this is particularly so for the more attractive ones (also known as

[16] Note, that in this and all subsequent tables we also condition on age and employment status.

power singles in the terminology of Costa & Kahn (2000) and Compton & Pollak (2007)). In fact, Compton & Pollak (2007) show that the reason that power couples are more often observed in cities (the Costa & Kahn, 2000 finding) is that marriage formation among power singles is more likely to happen in cities. Our analysis provides additional evidence that after marriage formation, power couples (here illustrated by the power man or woman who tends to marry power spouses (see e.g. Nielsen & Svarer, 2009)) leave the city at a higher rate and locate in less populated areas[17].

Table 4 also shows that the presence of children is associated with a lower mobility propensity between the city and the countryside and vice versa. This suggests that children raise moving costs. The marginal effect of children is higher for the migration from the countryside to the city, which indicates that amenities of rural living like more space, cheaper housing and less pollution are more valuable to individuals with children. In 5.2, we investigate further whether the presence of children is the driving force behind the mobility differences between singles and married individuals. For now, we simply notice that whereas children are affiliated with a lower mobility probability from the countryside to the city they are also associated with a lower probability of moving from the city to the countryside.

Finally, we give additional life cycle evidence in Table 5. We present the distribution of individuals between rural and city conditional on their level of attractiveness at different life stages.

Table 5: Geographical location of individuals based on attractiveness at different life-stages[18]

	Fraction living in city		
Level of attractiveness by quartile	At age 18	At marriage	After 5 years of marriage
1	0.22	0.28	0.20
2	0.20	0.27	0.18
3	0.23	0.33	0.22
4	0.18	0.50	0.32

[17] In the analysis we do not model the dynamic mobility process of each individual. As a short cut we treat each annual observation as independent. This approach could be invalidated if there is a lot of return migration. However, among the married couples that leave the city around 90% do not return to a city during their marriage. If we exclude the return migrants from the analysis our qualitatative findings remain the same.

[18] Since there are no differences in the gender specific patterns we present numbers for both men and women in this table.

Table 5 shows that at age 18 (when individuals are typically not yet active on the marriage market and often still live with their parents) most people live in rural areas and there is not a lot of difference between individuals at the high and the low end of the attractiveness distribution. However, at the time of marriage, many of them live in the city. In particular, a large share of the individuals in the fourth quartile have moved to the city and married there. Amongst the individuals who stay married for 5 years we see that a significant fraction has moved back to the rural areas and a pattern arises that is very close to the pattern at age 18. Only the most attractive individuals prefer to live in the city, although also for this group, we see transitions from city to rural upon marriage.

Compton & Pollak (2007) look at the proportion of power men and women living in US cities with more than 2 million inhabitants. They also find that conditional on education, singles are more likely to live in a large city than couples. This observation is consistent with the Danish data.

4.1 Size of the segments

In Section 2.3, we suggested that cities have more and smaller segments which implies a higher correlation between levels of attractiveness of men and women.

In this section we present evidence for this claim. We investigate whether differences in our attractiveness measure between partners differ between the city and rural areas. Below, we present the Spearman rank correlation between attractiveness and education separately.

Table 6: Spearman rank correlation at time of marriage

Spearman correlation	City	Rural
Attractiveness	0.404	0.286
Education	0.393	0.262
# Observations	4944	9823

For both measures of attractiveness the rank correlation is larger in the city which suggests a finer segmentation. The evidence presented above is to a large extent consistent with the marriage-market hypothesis presented in Section 2. In the subsequent sections we look at a number of alternative explanations. The main goal is to see whether the marriage market story still prevails once these confounding mechanisms have been included in the analysis.

5 Alternative explanations and robustness checks

In this section we carry out a number of sensitivity checks and test whether our results can be driven by other factors. First, we test whether the inflow of singles into the city merely reflects a "college effect", and second we test whether the fact that couples move out of the city is mainly due to the presence of children and finally we test whether our results could be driven by life cycle motives.

5.1 Going to the city to get a college education

In Denmark, most universities are located in the larger cities so we must worry about whether our results are driven by youngsters who move into the city to get an education, get married and then move back to the countryside. First, this story is not necessarily inconsistent with our marriage market model because colleges and universities are good marriage markets themselves because they select a fairly homogeneous group of highly educated individuals (see e.g. Goldin, 1992, Goldin and Katz, 2002, and Nielsen & Svarer, 2009). Nevertheless, it is useful to check whether our predictions still hold in the absence of colleges. We can do this by restricting the sample to individuals who are older than 25 years and assume that the motivation for those individuals to move to the city cannot be the presence of colleges.[19] The results of this exercise are presented in Table 8.

[19] This will be violated if a college graduate from Copenhagen who moved out after graduating is more likely to move back in because she knows her way around the city than a non-college graduate.

Table 8: linear probability model: individuals older than 25

	Women				Men			
	Rural to city		City to rural		Rural to city		City to rural	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Single	0.032*	0.002	-0.018*	0.006	0.021*	0.002	-0.009*	0.006
Attractiveness	0.002*	0.001	0.015*	0.002	0.005*	0.001	0.013*	0.002
Single*attract.	0.012*	0.002	-0.016*	0.005	0.010*	0.002	-0.002	0.004
Children	-0.013*	0.001	-0.013*	0.005	-0.014*	0.001	-0.005	0.004
# Observations	51764		17154		46528		16591	

Note: *denotes significantly different from 0 at the 5% level.
We condition on age and whether the individual works full-time.

Comparing the results in Table 8 to the results found for the unrestricted sample reveals that the single and attractive men and women are even more likely to move to the city than their married counterparts, and that this is especially so for the more attractive singles. Also the mobility patterns from the city to the countryside are consistent with the results found for the complete sample. Even though school attendance presumably is a major factor to locate in cities, those who are beyond the schooling age and are single are also strongly attracted to the cities.

5.2 The role of children

Although we control for having children, married couples could still move to rural areas because they expect to have kids in the future. In that case, the reason to move to the countryside reflects a shift towards more space and not the fact that one loses the benefits of a higher contact rate. In order to isolate the search motivation, we only consider the subset of couples who never get children. Under the assumption that having no kids reflects preferences rather than constraints, this group must have other motives than kids to move to the countryside.

Table 9: linear probability model: no kids sample

	Women				Men			
	Rural to city		City to rural		Rural to city		City to rural	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Single	0.024*	0.006	-0.018	0.012	0.011**	0.006	-0.011	0.012
Attractiveness	0.006*	0.003	0.003	0.005	0.011*	0.003	0.004	0.005
Single*attract.	0.010*	0.005	-0.002	0.008	-0.001	0.005	-0.008	0.009
# Observations	4926		2129		4915		2184	

Note: *(**) denotes significantly different from 0 at the 5% (10%) level.
We condition on age and whether the individual works full-time.

The absence of kids does not change the sign of the coefficients. We do, however, see that the standard errors increase and that in terms of statistical significance our results are less clear for the reduced sample. For the transition from the countryside to the city we still find that singles have a higher moving propensity and that this is especially pronounced for the more attractive single women. We also estimated mobility equations (not shown) for a sample of mothers. That is, we investigate whether single mothers are more likely to move than married mothers. We basically find the same pattern as in Table 8. Single mothers are more likely to move to the city, but not statistically more likely to stay in the city compared to married mothers. One interpretation of the results in Table 8 is that even if different preferences related to children are controlled for, singles are more likely to move to the cities due to the better marriage market conditions there. Although the signs are as predicted, our estimations are not strong enough to conclude that married individuals without kids are more likely to leave the city. A possible reason for this result is the higher mobility costs associated with having children. For future research it would be relevant to investigate this issue in more depth with a larger data set to get more precise estimates.

5.3 Life cycle motives for leaving the city

Although our theoretical model does not allow for divorce and return migration, both are salient facts in the data.[20] The behavior of divorcees can shed light on the geographical mobility of single people. Indeed, the mobility pattern that we reported so far could be the result of "ordinary" life cycle behavior. People enter the city when they are young and

[20] See for a dynamic model of return migration Kennan and Walker (2008).

have relatively strong preferences for bars, clubs, cinemas and other city amenities and then leave the city when they are older and richer and have stronger preferences for space. One way to "isolate" the search effect is to consider the mobility patterns of couples who have moved to the country and who divorced there[21]. If they moved to the countryside for life cycle motives other than the marriage market, we expect them to stay in the country after divorce whereas according to the marriage market model they should move back to the city once they become single again. We find that our model still holds. Since the observations we use are annual, we only know that a divorce has occurred during the year but not the exact date. We therefore present results for both individuals who have been divorced 1 year and those who are still divorced after 2 years.

Table 10: linear probability model: divorcees

	Women				Men			
	Rural to city		City to rural		Rural to city		City to rural	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Divorced, 1 year	0.051*	0.003	0.071*	0.009	0.026*	0.004	0.052*	0.009
Attractiveness	0.009*	0.001	0.007*	0.001	0.013*	0.001	0.009*	0.001
Divorced∗attract.	0.020*	0.004	-0.007	0.007	0.019*	0.004	-0.020*	0.007
Children	-0.026*	0.001	-0.003	0.004	-0.019*	0.004	0.002	0.004
Divorced, 2 years	0.019*	0.005	-0.021**	0.012	0.017*	0.005	-0.021**	0.012
Attractiveness	0.010*	0.001	0.007*	0.001	0.013*	0.001	0.009*	0.001
Divorced∗attract.	0.018*	0.005	-0.004	0.009	0.012*	0.006	-0.012	0.010
Children	-0.027*	0.001	-0.005	0.004	-0.019*	0.001	-0.001	0.004
# Observations	84736		29714		76130		27065	

Note: *(**) denotes significantly different from 0 at the 5% (10%) level.
We condition on age and whether the individual works full-time.

Not surprisingly, because of the nature of a divorce, divorcees are more likely to move. Therefore we must compare the likelihood to move to the city with the likelihood to move out of the city. For men, there is no large difference after the first year of divorce. In the second year after divorce they are however more likely to move to the city, but not to the rural areas compared to the reference group. For both men and women, the propensity of the divorced to move to the city is larger than to move out of the city after 2 years of divorce but not after 1 year of divorce. This can be explained by the fact that it is typically harder to find a place to live in the city than in the countryside.

[21] An ending cohabitation also counts as a divorce.

5.4 Other alternatives

A parallel explanation to our story is that people who live in the country can meet as many potential marriage market partners as people who live in the city but it is more costly for them to do so (i.e. they have to drive to a city, take a voluntary job where one meets many people etcetera). The opportunity cost of doing so is higher for more skilled workers so they opt for living in the city.[22] We could model this by introducing search cost to our model but this would add yet another source of multiplicity similar to Diamond (1982) and make it intractable.

Another possibility is that high skilled workers (which are the more attractive workers in our empirical application) must work harder in dense areas to distinguish them from rivals, see Rosenthal and Strange (2008). If the participants of the urban rat race are more likely to leave the city after a few years of 80 hour working weeks this could explain the patterns of table 5. However, it cannot explain why couples are more likely to leave the city than singles.

6 Final remarks

In this paper we extend the Burdett-Coles (1997) marriage market model with a distinction between efficient marriage markets (cities) and less efficient search markets (rural areas) and derive how individuals sort into those markets. Our model predicts that singles and in particular attractive singles move to the city while couples move out of the city. Those predictions are confirmed by the data. We find that in particular for females, the cross partial of single and attractiveness on the probability of moving is positive and statistically significant. Why the cross effect is less pronounced for males is still an open issue.

We offer a range of alternatives stories that can generate similar mobility patterns as those derived from our marriage market hypothesis. These include: higher returns to education in the city, the presence of universities in cities, and life cycle motives for moving in and out of the city. By using appropriately chosen sub samples, we show that there is still room for the marriage-market motivation. Finally, one might wonder why cities have more amenities per capita that are aimed at singles like bars and clubs? We

[22] We thank one of the referees for pointing out this possibility.

view this as the natural market response to the desire of singles to cluster in cities.

For future research we believe it could be fruitful to include divorce decisions. One interesting motivation for married couples to move to a rural area is that it is an efficient way to make a commitment. Burdett et al. (2004) showed that if one of the partners is likely to continue searching "on-the-job", this by itself stimulates the other partner to continue search as well. Given the many long term investments that are required, like raising children and buying a house, which all require a stable relationship, it can be efficient to move to an inefficient search market to limit "on-the-job" search. In Gautier, Svarer and Teulings (2008) we investigate this further. Furthermore, it would be interesting to allow for investments in one's type, as in Burdett, Coles (2001) and Mailath, Samuelson and Shaked (2000).[23] The latter show, also using an increasing returns to scale contact technology, that if workers are sorted on the basis of pay-off irrelevant characteristics (i.e. green and red) that there exist equilibria where firms spend more effort to search in the green-worker areas and the green workers invest more in human capital than the red workers. In our model, attractiveness is exogenous but their results imply that there may exist equilibria where attractive females mainly search in cities and where males invest more in their types. Finally, we focussed on the marriage market in this paper and abstracted away from the labor market. This does not mean that we believe that labor market considerations are not important to understand migration. To the contrary, we believe that an interesting avenue for future research is to jointly study the labor and the marriage market in one model. In such a model it would be important to consider joint location constraints and household bargaining as in Gemici (2008).

References

[1] Akerlof, G. (1985). "Discriminatory, Status-based Wages among Tradition-oriented, Stochastically Trading Coconut Producers", *Journal of Political Economy*, 93, (2), 265-76.

[2] Anderberg, D. (2004). "Getting Hitched: The Equilibrium Marriage Market Behaviour of a British Cohort", Manuscript, Royal Holloway University of London.

[23] See also Akerlof (1985) for a related model on discrimination with IRS.

[3] Aumann, R.J. (1959). "Acceptable Points in General Cooperative n-Person Games", in *Contributions to the Theory of Games IV*, A.W. Tucker and R.D. Luce, Eds., Princeton University Press, 287-324

[4] Bacolod, M., B. Blum, and W. C. Strange (2009), "Elements of Skill: Traits, Intelligences, Education, and Agglomeration", Working paper, Rotman School of Management Toronto.

[5] Becker, G (1973). "A Theory of Marriage: Part I", *Journal of Political Economy*, 82(4), 813-846.

[6] Beehler, B. and M.S. Foster (1988). "Hotspots, Hotshots and Female Preference in the Organization of Lek Mating Systems", *American Naturalis*, 131, 203-219.

[7] Bradburry, G (1981). "The Evolution of Leks", *Natural Selection and Social Behavior*, (R.D. Alexander and D. Tinkle, Eds.), 138-169, Chiron Press, New York.

[8] Bloch F. and H. Ryder (2000). "Two-Sided Search, Marriages, and Matchmakers", *International Economic Review*, 41(1), 93-115.

[9] Burdett, K. and M. G. Coles (1997). "Marriage and Class", *Quarterly Journal of Economics*, 112, 141-168.

[10] Burdett, K. and M. G. Coles (1999). "Long-term Partnership Formation: Marriage and Employment", *Economic Journal*, 109, 307-334.

[11] Burdett, K., R. Imai and R Wright (2004). "Unstable Relationships", *Frontiers of Macroeconomics*, 1 (1), Article 1.

[12] Burdett, K. and M. G. Coles (2001). "Transplants and Implants: The Economics of Self-Improvement", *International Economic Review*, 42, 597-616.

[13] Black, D., G. Gates, S. Senders and L. Taylor (2002). "Why Do Gay Men Live in San Francisco?", *Journal of Urban Economics*, 51, 54-76.

[14] Coles, M.G. and M. Francesconi (2008), On the Emergence of Toyboys: Equilibrium Matching With Ageing and Uncertain Careers, mimeo Essex.

[15] Combes, P.P., G. Duranton and L. Gobillon (2008), Spatial Wage Disparities: Sorting Matters!," *Journal of Urban Economics*, vol. 63(2), pages 723-742.

[16] Compton, J. and R. Pollak (2007). "Why are Power Couples Increasingly Concentrated in Large Metropolitan Areas?", *Journal of Labor Economics*, 25, 3, 475-512.

[17] Costa, D. L. and M. E. Kahn (2000). "Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990", *Quarterly Journal of Economics*, 1287-1315.

[18] Dahl M.S. and O. Sorenson (2008). "The Social Attachment to Place", Working Paper 08-34, University of Aalborg.

[19] Diamond P.A. (1982). "Wage Determination and Efficiency in Search Equilibrium", *Review of Economic Studies*, 217-227.

[20] Eeckhout, J. (1999). "Bilateral search and Vertical Heterogeneity", *International Economic Review*, 40, 869-87.

[21] Eeckhout, J. (2006). "Minorities and Endogenous Segregation", *Review of Economic Studies*, 73, 31-53.

[22] Edlund, L. (2005). "Sex and the City", *Scandinavian Journal of Economics,* 107 (1), 25-44.

[23] Gale, D. and L. Shapley (1962). "College Admission and the Stability of Marriage", *American Mathematical Monthly,* 69, 9-15.

[24] Gautier, P.A. M. Svarer and C.N. Teulings (2009). "Sin City", *Scandinavian Journal of Economics* 111(3), 439–456.

[25] Gautier, P.A. and C. Teulings (2009), "Search and the City", *Regional Science and Urban Economics*, 39-3, 251-65.

[26] Gautier, P.A. and C. Teulings (2003), "An Empirical Measure for Labor Market Density", *Review of Economics and Statistics*, 85- 4, 901-8.

[27] Gemici, A. (2008). "Family Migration and Labor Market Outcomes", Mimeo, University of Pennsylvania.

[28] Goldin, C. and L. Katz (2002). “The Power of the Pill: Oral Contraceptives and Women’s Career and Marriage Decisions”, *Journal of Political Economy,* 110, 730-770.

[29] Greenwood, M.J. (1997). “Internal Migration in Developed Countries”, Handbook of Population and Family Economics, ed. M. Rosenzweig and O. Stark, Vol. 1B, 647-720.

[30] Hamermesh, D. and J. Biddle (1994). “Beauty and the Labour Market”, *American Economic Review*, 84, 5, 1174-1194.

[31] Heckman, J. and J. Scheinkman (1987). "The Importance of Bundling in a Gorman-Lancaster Model of Earnings," *Review of Economic Studies*, 54(2), 243-55.

[32] Jacquet, N. L. and S. Tan (2007), On the Segmentation of Markets, *Journal of Political Economy*, 115, 639-664.

[33] Johnson, R.A. and D. W. Wichern (1998). Applied Multivariate Statistical Analysis, 4th edition, Prentice Hall, New Jersey.

[34] Mailath G.J., L. Samuelson, A. Shaked (2000). “Endogenous Inequality in Integrated Labor Markets with Two-Sided Search”, *American Economic Review*, 90, 46-72.

[35] Mincer, J. (1978). “Family Migration Decisions”, *Journal of Political Economy,* 86, 749-773.

[36] Nielsen, H. and M. Svarer (2009). “Educational Homogamy: How Much is Opportunities?”, *Journal of Human Resources,* 44(3), 1066-1086.

[37] OECD (2008). "International Migration Outlook", Paris.

[38] Rosenthal, S.S. and W.C. Strange (2004). Evidence on the Nature and Sources of Agglomeration Economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland.

[39] ROSENTHAL, S.S. AND W.C. STRANGE (2008), "Agglomeration and Hours Worked," *Review of Economics and Statistics* 90, 105-118.

[40] SANDER, W. (1985), "Women, Work and Divorce", *American Economic Review* 75, 519-523.

[41] SMITH, J.P. AND D. THOMAS (1998). "On the Road: Marriage and Mobility in Malaysia", *The Journal of Human Resources,* 33, 805-832.

[42] SMITH, L. (1997). "The Marriage Model with Search Frictions", mimeo University of Michigan.

[43] STARK, O. (1988). "On Marriage and Migration", *European Journal of Populations,* 4, 23-37.

[44] SVARER, M. (2004). "Is Your Love in Vain? Another Look at Premarital Cohabitation and Divorce", *Journal of Human Resources,* 39(2), 523-536.

[45] SVARER, M. AND M. VERNER (2008). "Do Children Stabilize Relationships in Denmark?", *Journal of Population Economics,* 21, 395-417.

[46] TEULINGS C.N. AND P.A. GAUTIER (2004). "The Right Man for the Job", *Review of Economic Studies,* 71(2), 553-580.

[47] WAGNER, R. (1997). "Hidden Leks: Sexual Selection and the Clustering of Avian Territories", *Ornithological Monographs*, 123-45.

[48] WONG, L. (2003). "Structural Estimation of Marriage Models", *Journal of Labor Economics*, 21(3), 699-727.

Appendix

A Derivations and omitted proofs

A.1 The derivation of equation $f_l(a)$

Integrating (4) over $m_l(a)$ and exploiting symmetry, $f_l(\alpha) = f_l(a)$, yields:

$$\int_{m_l(a)} g_l(\alpha)\, d\alpha = \lambda_l \left(\int_{m_l(\alpha)} f_l(\alpha)\, d\alpha \right)^2 + \delta \int_{m_l(\alpha)} f_l(\alpha)\, d\alpha \tag{15}$$

Solving (15) for $\int_{m_l(a)} f_l(\alpha)\, d\alpha$ and substitute this back in (4) yields the following expression for the steady state mass of females of type a:

$$f_l(a) = \frac{2g_l(a)}{\delta + \sqrt{\delta^2 + 4\lambda_l \int_{m_l(a)} g_l(\alpha)\, d\alpha}}. \tag{16}$$

A.2 The marriage market equilibrium in a single region

By equation (16), $f_l(a)$ satisfies

$$\begin{aligned} f_l(\alpha) &= f_l(a) = \frac{2g_l(a)}{\delta + \sqrt{\delta^2 + 4\lambda_l G_l\left(a_l^i, a_l^{i-1}\right)}}, \\ G_l\left(a_l^i, a_l^{i-1}\right) &\equiv \int_{a_l^i}^{a_l^{i-1}} g_l(\alpha)\, d\alpha. \end{aligned} \tag{17}$$

Define:

$$\begin{aligned} a_l^0 &\equiv a^+, \\ a_l\left(a, a_l^{i-1}\right) &\equiv \frac{1}{\delta}\sqrt{\frac{\lambda_l}{G_l\left(a, a_l^{i-1}\right)}} \int_a^{a_l^{i-1}} (\alpha - a)\, g_l(\alpha)\, d\alpha - lC. \end{aligned} \tag{18}$$

The last expression gives the utility of singles whose upperbound is a_l^{i-1} for various values of the lower bound a. In equilibrium, the lower bound of one's matching set is the value that makes a single indifferent between marriage and remaining single.

An equilibrium in the marriage market is a collection of connected non overlapping segments such that the lower bound a_l^i of each segment i at location l is the upper bound of the next segment. Then, the lower bounds a_l^i of segments i are determined recursively starting from $i = 1$ by the following algorithm:

$$a_l^i = a_l\left(a_l^i, a_l^{i-1}\right),$$

except for the lowest segment, denoted I_l, for which

$$\begin{aligned} a^- &> a_l\left(a^-, a_l^{I_l - 1}\right), \\ a_l^{I_l} &= a^-. \end{aligned}$$

Furthermore,

$$u_s(a, l) = a_l^i - lC, \forall a \in \langle a_l^i, a_l^{i-1}]. \tag{19}$$

A sufficient condition for uniqueness is:

$$2H(a,b)^2 > h(a)\int_a^b (\alpha - a)\, h(\alpha)\, d\alpha, \forall (a,b), a^- \leq a < b \leq a^+, \tag{20}$$
$$H(a,b) \equiv \int_a^b h(\alpha)\, d\alpha.$$

This can be seen as follows. The proposition generalizes the logic of equation (8) and (9). For $a = a_l^i$, equation (18) is identical to equation (9), where we substitute $f_l(\alpha)$ for equation (16). Given the upper bound a_l^{i-1} of a segment i, we can calculate $a_l\left(a, a_l^{i-1}\right)$ as a function of a.

Suppose condition (20) holds. Then, this condition also holds for $g_l(a)$ within a segment: if mobility cost are infinite, it holds by equation (7); if mobility cost for singles are zero, either all attractiveness levels within a segment prefer region l or none since all have the same utility as a single $u_s(a,l) = a_l^i - lC$ within a location, so either the one location dominates or the other, so that $g_l(a) = h(a)$ within that segment, see equation (6). For $a = a_l^{i-1}$, $a_l\left(a, a_l^{i-1}\right) = -lC < a$ and

$$\frac{\partial a_l\left(a, a_l^{i-1}\right)}{\partial a} = -\frac{1}{2\delta}\sqrt{\lambda_l G_l\left(a, a_l^{i-1}\right)}\left[2G_l\left(a, a_l^{i-1}\right)^2 - g_l(a)\int_a^{a_l^{i-1}} (\alpha - a)\, g_l(\alpha)\, d\alpha\right]$$
$$< 0, \forall a < a_l^{i-1}. \tag{21}$$

where the inequality follows from the fact that condition (20) applies. Hence, $a_l\left(a, a_l^{i-1}\right) = a$ has either a unique interior solution for $a < a_l^{i-1}$ or it has no solution at all, in which case i is the lowest segment I_l at location l, for which $a_l^i = a^-$.■

Without condition (20), we cannot rule out multiple equilibria. The source of the multiplicity of equilibria is the same as in Burdett and Coles (1997): if all women use a non-selective acceptance strategy, where they also marry with unattractive males, they all marry fast. By symmetry, this implies that males do the same and the stock of singles is small. This reduces the probability of finding a very attractive male, and makes applying a more selective strategy not profitable. However, if the most attractive females in the segment are all selective, then they stay single for a longer period. This raises the probability of finding an attractive partner, and hence the selective strategy is an equilibrium as well. Since this multiplicity is not central to our argument, we rule it out by imposing condition (20).[24]

[24] Condition (20) holds for example for the uniform distribution. The condition is sufficient but not

A.3 Proof of Proposition 2

It is sufficient to prove that if singles in segment i prefer the countryside above the city, so will singles in segment $i+1$. The location of segment $i-1$ is irrelevant, so we omit the suffix l for that segment. Suppose segment i prefers the countryside above the city $a_0^i < a_1^i$. By equation (6), this implies

$$\begin{aligned} a_0^i &= \frac{1}{\delta}\sqrt{\lambda_0}\int_{a_0^i}^{a^{i-1}} \left(\alpha - a_0^i\right) h(\alpha)\, d\alpha > \frac{1}{\delta}\sqrt{\lambda_1}\int_{a_1^i}^{a^{i-1}} \left(\alpha - a_1^i\right) h(\alpha)\, d\alpha - \delta C \\ &> \frac{1}{\delta}\sqrt{\lambda_1}\int_{a_0^i}^{a^{i-1}} \left(\alpha - a_0^i\right) h(\alpha)\, d\alpha - \delta C. \end{aligned}$$

The second inequality follows from (21). The first equality implies

$$\frac{\delta a_0^i}{\sqrt{\lambda_0}} = \int_{a_0^i}^{a^{i-1}} \left(\alpha - a_0^i\right) h(\alpha)\, d\alpha.$$

Substitution in the second inequality yields:

$$\frac{\sqrt{\lambda_1} - \sqrt{\lambda_0}}{\sqrt{\lambda_0}} a_0^i < \delta C.$$

Since $a_0^i > a_0^{i+1}$, this inequality holds also for segment $i+1$. Clearly, for low values of C, the final condition is not satisfied for any segment; for high values of C, it is satisfied for all segments.■

necessary. A stricter condition can be derived from the inequality: $\partial a_l\left(a, a_l^{i-1}\right)/\partial a < 1$.All subsequent arguments hold in the case of multiplicity if we add a criterium for equilibrium selection that requires all types above the attractiveness level a_l^i to be able to solve their coordination problem by following the selective strategy. This strategy maximizes their expected utility. However, this only adds to the complexity of the argument. Finally, note that Burdett and Coles (1997) derived that in their setting log concavity of the type distribution is sufficient for uniqueness.