Sin City: Why is the divorce rate higher in urban areas?*

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

Abstract

We show with Danish data that, of the couples who married in the city, the ones who stay in the city have a 23% higher divorce rate than the ones who move to non-urban areas. Similarly, for the couples who married outside the city, the ones who move to the city are more likely to divorce. This correlation can be explained by both a causal and a sorting effect. We disentangle them by using the timing-of-events approach. In addition we use information on father's location as an instrument. We find that after allowing for sorting, the effect of living in the countryside on the divorce hazard drops substantially and looses statistical significance.

Keywords: Dissolution, search, mobility, city.

Classification-JEL: J12, J64

*Acknowledgement: Michael Svarer acknowledges financial support from the Danish National Research Foundation through its grant to CAM and the Danish Social Science Research Council. We are grateful to the editor Nils Gottfries and two referees for constructive and useful comments. We further thank Jaap Abbring and seminar participants at Cemfi Madrid, CERGE Prague, Copenhagen Business School, Centre for Applied Microeconometrics Copenhagen, University of Aix-Marseille 1, Tinbergen Institute Amsterdam, and IZA, Bonn for their comments. We finally thank Ulla Nørskov Nielsen for very useful research assistance.

1 Introduction

We give evidence that of the marriages that are formed in the city, those couples who remain in the city have a higher divorce rate than the ones who move out. Likewise, the couples who marry in the countryside but move to the city are more likely to divorce than the ones who stay in the countryside. The main question we want to address in this paper is whether this correlation reflects a causal link. In Gautier et al. (2005) we give evidence that cities serve as a marriage market. The basic idea is that the rate at which singles meet potential partners is higher in the city either because of a size-of-the-market effect or because cities are more densely populated. Therefore, singles (in particular the most attractive ones) will exploit this and move to the city.[1] The same observation suggests that leaving the city will stabilize relationships. By moving to the countryside, the number of outside offers decreases for both partners which on its turn decreases the value of continued search while married. This is consistent with the fact that couples have a larger probability than singles to leave the city, even those who never have kids. Alternatively, if relatively unstable relationships sort themselves in the city then we also observe a higher divorce rate in cities but then there is no causal link. One possible story that is consistent with sorting is that stable marriages are more likely to want kids and are more likely to buy a house, see Svarer and Verner (2008) for evidence. Since kids require more space and since there is more home ownership outside the city we find a large proportion of the stable marriages outside the city. Another possibility is that living in a remote area is attractive because of the low land prices but it also implies that one has to spend a large fraction of time together with one's partner and this requires a stable relationship.

There is some evidence that exposure to easier divorce regulation when young worsens adult outcomes along a number of dimensions including education, see Gruber (2000). If urbanization destabilizes marriages and if divorces are bad for kids then policy makers

[1]Comton and Pollak (2007) argue that ammendities do not follow singles but that singles follow ammendities. The city ammendities attract singles and therefore marriage formation is more likely to occur in cities. Costa and Kahn (2000) argue that power couples are more likely to go or stay in the city because of the double job search problem. Finally, Edlund (2005) argues that young women outnumber young men in urban areas. The argument is that urban areas offer skilled workers better labor markets and the presence of high skilled males attracts females of all skills..

should take this into account when they make plans to either stimulate or discourage urbanization. On the other hand, Piketty (2003) finds that children of a couple that divorces perform equally bad in the years before their parents separate. Björklund and Sundstrom find that siblings who were not personally exposed to the separation of their parents have an educational attainment that is the same as that of their younger brothers and sisters that were exposed to separation. So the evidence on whether divorces have a causal effect on children's outcomes or whether everything is driven by marriage quality is inconclusive.

Another potential reason why the divorce rate could be inefficiently high is given by Burdett et al. (2004). They show that even good marriages can become unstable if agents believe that their partners continue to search for a new partner. In that case, the best response is to also start searching. They give an example where a faithful equilibrium where nobody searches while being married is the most efficient one but where the market selects the inefficient unfaithful equilibrium where both partners continue search. If search cost are sufficiently high, the unfaithful equilibrium no longer exists which is welfare improving for the agents. If there exists a causal relation between the higher arrival rates of new marriage partners in the city and the higher divorce rates than couples could benefit from leaving urban areas to areas with a lower arrival rate where only the faithful equilibrium exists.

If the outcome of the aging process is a random variable this can potentially also destabilize marriages if for one of the partners the outcome of this process is more favorable than for the other, see Masters (2006). He suggests that the more attractive aging partner voluntarily becomes less attractive (i.e. by increasing weight) in order to stabilize the relationship. This is however a costly commitment relative to moving to the countryside.

In the sociological literature, there has been a long tradition of relating urbanization and family change (including divorces).[2] To our knowledge this literature has not established a convincing causal link between cities and divorce rates. Gemici (2006) also considers migration decisions for married couples and singles but she focuses more on labor market returns. Chiappori et al. (2005) consider a marriage market with transferable utility. Since there is continuous renegotiation possible, inefficient separations do not

[2] This literature goes back to Calhoun (1945) and Burgess et al. (1963). Schultz cites a special report from the Census (1909) which demonstrates that the urban counties in the North Central region had consistently higher divorce rates than the less urbanized counties.

occur.

Higher divorce rates in urban areas do of course not imply a causal effect. There could for example exist a correlation between preferences for living in the country side and the quality of a marriage. I.e. stable relationships are more likely to move to remote areas. Therefore, a lower divorce rate in rural areas can just reflect a correlation of unobservable characteristics and location choice.

We apply the timing-of-event methodology (Abbring and van den Berg, 2003) to distinguish the causal effect of living in a city on the divorce rate from the correlation-through-unobservables effect. In addition, we conduct an instrumental variable analysis using information on father's location as exclusionary restriction. The assumption is that father's location affects moving decisions but not the stability of a marriage. Our results suggest that the sorting effect is important and that there is no statistically significant causal effect of living in the city on divorce rates.

The paper is organized as follows, in section 2 we discuss the data and our empirical strategy. Section 3 contains our empirical results and section 4 concludes.

2 Data and empirical model

The data that we use to test the main implications of the model come from IDA (Integrated Database for Labour Market Research) created by Statistics Denmark. The information comes from various administrative registers that are merged in Statistics Denmark. The IDA sample used here contains (among other things) information on marriage market conditions for a randomly drawn sub-sample of all individuals born between January 1, 1955 and January 1, 1965. 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 observe an individual's mobility pattern on an annual basis. If the individual enters a relationship we also observe the personal characteristics of the partner. Based on the available information we sample all partnerships that are formed during the observation period. That is, we follow much of the duration literature (see e.g. van den Berg (2001)) and base inference on a flow sample of partnership by discarding those partnerships that were formed before 1980. With respect to the movement process we set the clock at zero at the moment of marriage.

We divide Denmark into two regions: cities and rural areas. In the main part of this paper we only include Copenhagen, the most dense area in Denmark which hosts 12.7 % of the population in 1995, in the city category and the rest of Denmark is considered to be the countryside. We also experiment with different city definitions but this does not change our conclusions. The main explanatory variable in our analysis is thus an indicator variable that takes the value 1 if the individual is currently living in Copenhagen.

Individuals can occupy one of three states in the marriage market: single, cohabiting, or married. Cohabitation as either a prelude to or a substitute of marriage is very common in Denmark (see e.g. Svarer, 2004).[3] There are some caveats to this definition of marriage.[4] Some of the couples - presumably a small minority - who are registered as cohabiting are simply sharing a housing unit, and do not live together as a married couple. In this paper we combine the states of cohabitation and marriage.

2.1 Descriptive statistics

In the period from 2000-2004 on average around 14 marriages per 1000 existing marriage divorced during a year. The geographical diversity in divorce rates is relatively marked. In Copenhagen 20 out of 1000 existing marriages divorced, whereas only 10 out of 1000 existing marriages divorced in rural areas of Jutland (Source: Statistics Denmark ((19xx).

The moving patterns in Denmark go in the direction of a higher level of urbanization. Since 1981 the larger cities have experienced increasing population rates of around 0.3% annually, whereas urban areas have experienced a decline of -0.24% annually. (Source: Christoffersen, H. (2003)). Combining the two trends it can be expected that aggregate divorce rates in Denmark should increase if there is a causal impact on living in more populated areas on divorce risks.

Our main variable of interest is the city dummy. In addition, we also include three

[3]In the current data set around 78% of the couples who marry lived together before marriage. The occurance of cohabitation is also increasing substantially in other countries. In the U.S. the number of cohabiting couples has increased from 1.1 million in 1977 to 4.9 million in 1997 (see e.g. Svarer (2004) for more details on the development in cohabitation in the Western world).

[4]The sociological literature (see. e.g. Bennett et al. (1988) and Forste (2002)) suggests that lack of permanence and commitment between partners are primary features distinguishing cohabitation from marriage.

other important variables that are likely to affect marriage duration in the analysis. First, we distinguish between couples who are formally married or not by the indicator variable, *marriage*. Second, we consider the housing status of the couple in the sense that we discriminate between *home owners* and those who do not own their own house. Finally, we have an indicator variable, *children 0-6*, for the presence of children between 0 and 6 years old in the household. In Table 1 we present the association measure gamma.[5]

	Countryside	Married	Kids 0-6 years old	Homeowner
Countryside	1	0.57*	0.48*	0.71*
Married		1	0.57*	0.47*
Kids 0-6 yrs old			1	0.34*

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

Table 1: Association measure (Gamma) of the key variables

As Table 1 shows, the association between the four variables suggest they are strong complements. As Drewianka (2005) argues, this is not surprising since each of these features increases the relative value of a relationship and stimulates investments in the others.

We also include a number of additional explanatory variables in the subsequent analysis like dummies for educational attainment. The descriptive statistics can be found in Table 2. Some individuals may still be studying (we observe the current education at the time of observation). The educational variables are therefore also allowed to be time-varying. The reference group has less than high school education. Vocational education refers to individuals that have some sort of practical training, like carpenters etc. The other categories refers to different levels of further education. "Short" represents people who have studied for 14 years, "medium" stands for 16 years of education and "long" for at least 18 years. Next, we use information on gross income. Gross income is measured

[5] Gamma is calculated as $\frac{P-Q}{P+Q}$, where P is the number of pairs of the two indicator variables that take the same value (1 and 1 or -1 and -1) whereas Q is the number of pairs that takes opposing values (-1 and 1 or 1 and -1).

in 1980 prices and includes both labour and non labour income as well as received unemployment insurance benefits. We also include variables measuring the age of the partners as well as their age difference. The variable, *sickness*, is an indicator variable taking the value 1 if the individual receives sickness benefits during the year. As a general rule sickness benefits are received if a person has a spell of illness for more than 13 weeks. Each individual's degree of unemployment during the year is defined as the number of hours of unemployment divided by the number of potential supplied working hours. Finally, we have an indicator variable that takes the value 1 if the father (data limitations imply that we only observe location of father, not the mother, and we can also not see if they are still together) of at least one of the individuals in a given couple is living in the countryside. This variable works as an exclusionary restriction in the subsequent analysis where we explicitly model the moving decision from the city to the countryside and vice versa. Our conjecture is that having a father currently living in the countryside can have a pull effect on one's location decision but is unrelated to the quality of the marriage.

Couples formed in	city		countryside	
	Mean	Std. dev.	Mean	Std. dev.
City	1		0	
Married	0.13		0.11	
Homeowner	0.15		0.19	
Kids. 0-6 years old	0.06		0.08	
Kids. 7-17 years old	0.04		0.06	
Father living in countryside	0.65		0.96	
Male's education				
Vocational	0.33		0.52	
Short	0.06		0.05	
Medium	0.10		0.06	
Long	0.14		0.04	
Same level of education	0.70		0.50	
Male more educated	0.15		0.25	
Income (in dkk 1980 level)				
Female income	64,975	(39,435)	56,480	(35,743)
Male income	85,482	(62,555)	87,516	(52,924)
Age				
Female between 15-20	0.54		0.66	
Female between 21-25	0.33		0.23	
Female between 26-30	0.10		0.09	
Male between 15-20	0.38		0.48	
Male between 21-25	0.37		0.31	
Male between 26-30	0.18		0.14	
Female more than 4 years older	0.08		0.06	
Male more than 4 years older	0.25		0.26	
Sickness and unemployment				
Sickness, female	0.08		0.10	
Sickness, male	0.08		0.11	
Unemployment degree, female	0.09	(0.20)	0.13	(0.24)
Unemployment degree, male	0.09	(0.21)	0.11	(0.21)
Relationship duration (in years)	5.86	(3.92)	6.82	(4.3)
Fraction of couples who leave*	0.38		0.04	
Number of observations	3292		16646	

note:* denotes the fraction of couples formed in city (countryside) that move to the countryside (city)

Table 2: Descriptive Statistics (at the beginning of relationship)

Of the couples formed in Copenhagen around 38% move to a less populated area during the course of the relationship. On the other hand, only 4% of the partnerships that are formed in the countryside move to Copenhagen.

2.2 Empirical Model

In order to investigate the effect of locating in a given area on the dissolution risk we estimate a duration model where the random variable is the time spent in a given relationship. Since the location decision is potentially endogenous to the dissolution risk, our goal is to disentangle the causal effect of cities from the sorting effect. First, we apply the timing-of-event model of Abbring and van den Berg (2003). We estimate the process of dissolution simultaneously with the process of moving to a less populated area allowing the two processes to be interdependent through the error structure. Second, we use an exclusionary restriction to strengthen identification.

2.2.1 Timing-of-events method

The timing-of-events method enables us to identify the causal effect of location choice on the dissolution rate under some well-defined assumptions which we discuss below. The estimation strategy requires simultaneous modelling of the divorce rate and the moving hazard. Let $T_{r(elationship)}$ and $T_{m(ove)}$ denote the duration of a relationships and the duration till the agent moves in or out of a city (where the clock starts when the couple is married). Both are continuous nonnegative random variables. We allow T_r and T_m to interact through correlation of unobservables or through a possible treatment effect of moving in or out of the city. Suppose that each period, the couple draws an $r =$(utility in city/ utility in countryside), where r depends on for example job market opportunities. Let the marriage quality be given by $q \in [0, 1]$. Then, the optimal strategy is to define a reservation value $r^*(q)$ above which the couple moves to the city. Then T_m depends on the quality of marriage but not in a deterministic way. This randomness is necessary for identification. We assume further that all individual differences in the joint distribution of the processes can be characterized by observed explanatory variables, x, and unobserved variables, v. The moving incidence and the exit rate out of marriage are characterized by the moments at which they occur, and we are interested in the effect of the realization of T_m on the distribution of T_r. The distributions of the random variables are expressed in terms of their hazard rates $h_m(t|x_{m,t}, v_m)$ and $h_r(t|t_m, x_{r,t}, v_r)$. Conditional on x and v, we can therefore ascertain that the realization of T_m affects the shape of the hazard of T_r from t_m onwards in a deterministic way. This independence assumption implies that the causal effect is captured by the effect of t_m on $h_r(t|t_m, x_{r,t}, v_r)$ for $t > t_m$. This rules out that t_m affects $h_r(t|t_m, x_{r,t}, v_r)$ for $t \leq t_m$, i.e. anticipation of the move has no effect on the relationship hazard. This assumption will be falsified if one or both partners stop searching in the anticipation period before moving to the city or searches extra hard in the anticipation period before moving to the countryside. However, we justify the use of the model by the fact that the time span between the moment at which the anticipation occurs and the moment that the actual move takes place is relatively short compared with the duration of a marriage (the average duration of relationships is approximately 6.7 years in our sample while the average time to find a house is only a few months). This implies that the potential bias from anticipation is small.

Given the independence and no anticipation assumptions, the causal effect of moving on the divorce rate is identified by a mixed proportional hazard model. That is, it is a product of a function of time spent in the given state (the baseline hazard), a function of observed time-varying characteristics, x_t, and a function of unobserved characteristics, v

$$h\left(t|x_t, v\right) = \lambda\left(t\right) \cdot \varphi\left(x_t, v\right),$$

where $\lambda\left(t\right)$ specified as $exp(\lambda_m(t))$ is the baseline hazard and $\varphi\left(x_t, v\right)$ is the scaling function specified as $exp(\beta' x_t + v)$. More specifically the system of equations is:

$$h_m(t|x_{m,t}, v_m) = \exp(\beta_m' x_{m,t} + \lambda_m(t) + v_m) \tag{1}$$

$$h_r(t|t_m, x_{r,t}, v_r) = \exp(\beta_r' x_{r,t} + \delta D(t_m) + \lambda_r(t) + v_r),$$

where $D(t_m)$ is a time-varying indicator variable taking the value 0 before the couple moves, and 1 after the couple moves.

Intuitively, the timing-of-events method uses variation in marriage duration and in duration until moving (conditional on observed characteristics) to identify the unobserved heterogeneity distribution. The selection or sorting effect is captured by a positive correlation between v_r and v_m while the causal effect of living in the city on marriage duration is captured by the effect of the presence outside the city conditional on the observables and v_r and v_m. If couples who move to the city divorce fast, irrespective of how long they lived outside the city there is a causal effect of living in the city on the divorce rate. Alternatively, if only the couples who move to the city just after marriage divorce faster while the ones who move later do not divorce faster, there is a sorting effect. The most stable relationships are more likely to remain in the countryside for a long time because they are more likely to have kids or prefer to spend lots of time together while the relatively unstable relationships move to the city fast. This requires however that there is no interaction between marriage quality and treatment. If for example living in the city causes bad marriages to dissolve faster this also implies a positive correlation between v_r and v_m. In that case, if we would randomly pick a treatment group of 1000 couples from the countryside and place them in the city we would find a positive treatment effect caused by the unstable relations who divorce faster in the city than in the countryside. Abbring and Van den Berg (2003) show that under further proportionality assumptions a cross effect of marriage quality and the treatments (city and countryside) is identified by allowing the unobserved characteristics of the marriage quality ν_r to be different for the movers and the non movers. The time varying piecewise constant duration effect is then informative on the city effect. We do not travel this avenue because the assumption of independence between observables and non-observables after the move cannot be justified.

Alternatively, we impose an exclusionary restriction in the moving equation (this identification strategy is along the lines of e.g. Lillard (1993)). Specifically, we include as an extra explanatory variable in the moving hazard -an indicator variable that takes the value 1 if the father of one of the individuals in a given couple currently lives in the countryside- assuming that this variable does not affect the dissolution risk but it does affect the location of the couple. For the Likelihood function we refer to the appendix.

3 Results

Since the quality of a relationship may depend on the location of the marriage -agents who met in a city could have been more choosy because of the higher contact rate - we report the results separately for the subset of relationships that are formed in the city and those that are formed in the countryside.

The variables of interest are: being married, whether one owns a house, having young kids, and having older kids. The latter distinction is important because the cost of divorce is larger for when you have young kids. Of particular interest in this study is the time-varying indicator variable that denotes whether the couple is currently living in the city or in the countryside. In addition to this variable we also condition on the usual suspects in the divorce literature (see e.g. Svarer (2004)). In Table 3 and 5 below, we present three sets of results for partnerships that were initiated in the city and the countryside, respectively. The full sets of estimates are given by 4 and 3. First, we show the results for a model where we do not model the moving decision (model 1). Second, we take the moving decision into account and use the timing-of-event model to address the potential endogeneity of moving in relation to the dissolution risk (model 2). Third, we use as exclusionary restriction an indicator variable that takes the value 1 if the father of one of the individuals in a given couple lives in the countryside and 0 otherwise (model 3). Specifically, we include this variable in the moving hazard equation.

	Model 1		Model 2		Model 3	
			Timing-of-event		T-o-E and IV	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Countryside	-0.264*	0.077	-0.121	0.136	-0.146	0.135
Married	-1.499*	0.113	-1.474*	0.113	-1.499*	0.114
Kids 0-6 yrs old	-0.325*	0.070	-0.319*	0.071	-0.320*	0.071
Kids 7-17 yrs old	0.013	0.102	-0.009	0.101	-0.008	0.102
Homeowner	-0.233*	0.084	-0.288*	0.088	-0.293*	0.089
Father living in countryside**					0.660*	0.090
Corr(v_m, v_r)***			-0.209*	0.073	-0.190*	0.070
# couples		3292		3292		3292
Log likelihood		-4409		-7737		-7713

Note: * denotes significant different from 0 at the 5% level. ** gives the results from the moving hazard. ***The standard error for the correlation coefficient has been calculated based on 1,000 drawings from the multivariate normal distribution with matrix set equal to the estimated parameter vector and covariance matrix.

Table 3: Results for dissolution hazard for relationships formed in the city

Table 3 shows that the variables that increase the relative value of the relationship indeed decrease the divorce hazard significantly. Couples that leave Copenhagen experience a drop in the dissolution hazard of 23% (exp(-0.264)-1). To get an idea of the order of magnitude of those effects, the mean marriage duration for those who stay in the city is 7.9 years so a 23% drop in the divorce hazard is substantial.

The effect of the other variables coincides with previous research on partnership dissolution. Böheim and Ermisch (2001) also find that formally married couples are less likely to divorce than their cohabiting counterparts. Weiss and Willis (1997) and Peters (1986) among others find that children (especially when they are young) are associated with lower dissolution risk. Sullivan (1995) and Jalovaara (2001) also find that homeowners are less likely to divorce. Home ownership is however also a proxy for wealth which, like income stabilizes marriage. We have on the other hand not been able to locate any previous work that explores the effect of moving from more populated areas to lesser populated on the dissolution risk. Although the fact that the divorce risk is lower in rural areas has also been observed by Peters (1986) and Jalovaara (2001).

The results presented in model 2 suggest that moving to the countryside is not an exogenous event in relation to the dissolution process. Indeed, the significant effect of leaving Copenhagen vanishes once we model the moving decision simultaneously with the dissolution process. The coefficient on the mobility dummy is still positive and rather large, but so are the standard errors. Hence, the lack of precision might be responsible for the insignificant result portrayed. Taken at face value, model 2 suggests that based on unobservable factors, the stable relationships are more likely to leave Copenhagen and this association is responsible for about half of the reduction in the divorce hazard in model 1. The sorting effect is captured by the correlation between the unobserved heterogeneity terms in the moving hazard and the dissolution hazard. This correlation is significantly negative. As we argued before, this could also be caused by an interaction between marriage quality and living in the city (cities have a causal effect on dissolution of bad marriages). Model 3, where we introduce an exclusionary restriction in the moving equation suggests however that this is not the case. Couples where at least one partner has a father currently living in the countryside have a much higher moving probability. In fact, the hazard rate out of Copenhagen is 93% higher for these couples. Assuming that father's location is unrelated to marriage stability and assuming that the effect of location choice on father's location is independent of marriage quality this variable randomizes locations of couples (irrespective of marriage quality). With this exclusion restriction the effect of living in the countryside on the divorce hazard is the same as in model 2. Moreover, we show in Table 4 that the moving hazard is higher for couples who also invest in a house, having young children or who are formally married. This suggests that also in terms of observables, the stable relationships are more likely to move to the countryside. Finally, there is a positive correlation between divorce and unemployment.

	Timing-of-event model				Timing-of-event model with exclusionary restriction			
	Moving		Dissolution		Moving		Dissolution	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E	Coeff.	S.E
Homeowner	2.005*	0.082	-0.288*	0.089	2.157*	0.090	-0.293*	0.089
Married	0.469*	0.088	-1.474*	0.113	0.456*	0.091	-1.499*	0.114
Have moved from Copenhagen			-0.121	0.136			-0.146	0.135
Children 0-6	0.472*	0.076	-0.319*	0.071	0.469*	0.075	-0.320*	0.071
Children 7-17	0.209	0.136	-0.010	0.102	0.214	0.138	-0.009	0.102
Exclusionary restriction								
Father living in countryside					0.660*	0.090		
Vocational education, male	0.240*	0.096	-0.227*	0.090	0.281*	0.097	-0.227*	0.091
Short cycle further edu., male	0.351*	0.170	-0.320*	0.158	0.414*	0.178	-0.345*	0.160
Medium cycle further edu., male	0.277	0.147	-0.440*	0.132	0.242	0.147	-0.453*	0.133
Long cycle further edu., male	-0.148	0.122	-0.084	0.108	-0.198	0.123	-0.080	0.109
Couple have same level of edu.	0.063	0.091	-0.001	0.078	0.106	0.091	-0.002	0.079
Male more educated	-0.198	0.129	0.232*	0.112	-0.128	0.130	0.235*	0.113
Relationship number	-0.104*	0.063	0.245*	0.055	-0.071	0.065	0.248*	0.056
Female between 15-20	-0.313	0.271	0.003	0.209	-0.028	0.274	-0.045	0.210
Female between 21-25	-0.046	0.236	0.023	0.179	0.165	0.239	-0.015	0.180
Female between 26-30	-0.123	0.213	-0.155	0.157	0.036	0.214	-0.180	0.158
Male between 15-20	0.475*	0.221	0.140	0.181	0.584*	0.225	0.153	0.182
Male between 21-25	0.394*	0.177	-0.115	0.146	0.459*	0.180	-0.101	0.147
Male between 26-30	0.209	0.151	-0.052	0.120	0.230	0.153	-0.042	0.121
Female more than 4 years older	-0.698*	0.217	0.537*	0.178	-0.583*	0.226	0.514*	0.179
Male more than 4 years older	0.174	0.129	0.340*	0.110	0.180	0.130	0.346*	0.111
Female income	0.309*	0.098	-0.232*	0.064	0.339*	0.044	-0.229*	0.064
Male income	0.340*	0.047	-0.289*	0.096	0.285*	0.098	-0.298*	0.097
Sickness, female	-0.045	0.103	0.123	0.086	-0.041	0.104	0.122	0.086
Sickness, male	-0.065	0.124	-0.025	0.101	-0.062	0.124	-0.015	0.102
Unemployment degree, female	0.629*	0.181	0.140	0.132	0.660*	0.180	0.125	0.133
Unemployment degree, male	-0.195	0.195	0.469	0.135	-0.228	0.195	0.464	0.135
Mass points (v_m^2, v_r^2)	-3.026*	0.154	-2.227*	0.172	-3.031*	0.153	-2.276*	0.170
$p_1(v_r^1, v_m^1)$	0.173*	0.034			0.159*	0.662		
$p_2(v_r^2, v_m^1)$	0.330*	0.029			0.316*	0.145		
$p_3(v_r^1, v_m^2)$	0.274*	0.030			0.278	0.231		
$p_4(v_r^2, v_m^2)$	0.223*	0.026			0.244*	0.123		
$\mathrm{Corr}(v_r, v_m)$	-0.210*	0.073			-0.190*	0.070		

Note: * denotes significance at the 5 % level.The standard error for the correlation coefficient and probabilities has been calculated based on 1,000 drawings from the multivariate normal distribution with mean and covariance matrix set equal to the estimated parameter vector and covariance matrix.

Table 4: Full set of Results for Partnerships formed in Copenhagen

The process of moving to a new location is a stressful event. To what extent does this affect our divorce rate? If we assume that the process of moving can only have an effect on the hazard rate in the first 2 years we can control for it by allowing for (piecewise constant) time varying treatment effects of the moving variable in the dissolution hazard. We do however not find significant time varying effects of moving on the dissolution hazard and conclude that this exercise does not change our main findings presented in Table 3.

We do not consider the exogeneity status of kids and home ownership in this study. In a related study, Svarer and Verner (2008) take a closer look at children. They find that couples that are less prone to end their relationship are more likely to get children. Since couples with children are more likely to leave the city and are also more likely to buy a house this suggests indeed that stable relationships are more likely to have kids.

In Table 5 we consider the dissolution hazard of individuals who married in the countryside.

	Model 1		Model 2		Model 3	
			Timing-of-event		T-o-E and IV	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
City	0.298*	0.072	0.187*	0.077	-0.017	0.099
Married	-1.482*	0.052	-1.134*	0.040	-1.488*	0.052
Kids 0-6 yrs old	-0.202*	0.030	-0.199*	0.026	-0.201*	0.029
Kids 7-17 yrs old	0.213*	0.042	0.189*	0.035	0.203*	0.042
Homeowner	-0.381*	0.035	-0.302*	0.029	-0.388*	0.034
Father living in countryside**					-1.937*	0.175
Corr(v_r, v_m)***			0.413	0.283	0.221*	0.086
# couples		19938		19938		19938
Log likelihood		-22227		-26166		-26451

Note: * denotes significant different from 0 at the 5% level. ** gives the results from the moving hazard***The standard error for the correlation coefficient has been calculated based on 1,000 drawings from the multivariate normal distribution with matrix set equal to the estimated parameter vector and covariance matrix.

Table 5: Results for dissolution hazard for relationships formed in the countryside

Again, we find a positive association between living in a more populated area and the risk of dissolution which is of the same order of magnitude as for the couples formed in the city. The mean marriage duration for couples formed in non- urban areas and who stay there is 8.27 years. The association between divorce rate and city looses power once we address endogeneity with only the timing-of-event model but remains significant. The correlation between the unobserved heterogeneity terms is also insignificant for model 2. However, when we use our exclusion restriction we find that there is a large and significant positive association between the unobservables in the moving and dissolution hazard and that the effect of moving to Copenhagen is driven by this association and not by a causal effect generated by our proposed mechanism. Our interpretation of the difference between model 2 and model 3 is that identification gets stronger when we include the instrument. In model 2 identification in only based on the rather low fraction of couples that move (around 4% cf. Table 2). Apparently, this is not sufficient to identify the correlation between the two sets of unobservables and the sorting effect is not significant (as shown in Table 5). Identification improves with the inclusion of the very significant dummy variable for whether at least one of the fathers of the couple lives in the countryside. Finally, Table 3 summarizes the other coefficient estimates for the couples who formed in the the countryside. Again, we find that the probability to move and or divorce is lower for couples who also invest in a house, having young children or who are formally married. We also found the same relationship between divorce and income and labor market state as for the couples who formed in the city.

	Timing-of-event model				Timing-of-event model with exclusionary restriction			
	Moving		Dissolution		Moving		Dissolution	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E	Coeff.	S.E
Commitment variables								
Homeowner	-0.518*	0.076	-0.303*	0.029	-0542*	0.083	-0.389*	0.034
Married	-0.135	0.085	-0.134*	0.040	-0.188*	0.090	-1.488*	0.052
Have moved to Copenhagen			0.187*	0.077			-0.018	0.099
Children 0-6 years	-0.645*	0.098	-0.199*	0.026	-0.688*	0.102	-0.209*	0.030
Children 7-17 years	-0.794*	0.190	0.190*	0.035	-0.901*	0.195	0.203*	0.042
Exclusionary restriction								
Father living in countryside					-1.937*	0.174		
Other variables								
Relationship number	-0.096	0.061	0.259*	0.022	-0.068	0.065	0.371*	0.030
Female between 15-20	0.179	0.319	0.117	0.084	0.401	0.334	0.029	0.099
Female between 21-25	0.091	0.300	-0.031	0.072	0.297	0.312	-0.064	0.085
Female between 26-30	0.110	0.290	-0.110*	0.065	0.315	0.301	-0.154	0.074
Male between 15-20	0.251	0.216	0.055	0.076	0.374	0.228	0.111	0.087
Male between 21-25	0.125	0.185	-0.005	0.060	0.274	0.193	0.084	0.070
Male between 26-30	0.052	0.173	-0.001	0.051	0.183	0.178	0.073	0.058
Female more than 4 years older	-0.063	0.196	0.500*	0.068	-0.047	0.213	0.657*	0.087
Male more than 4 years older	-0.233	0.110	0.187*	0.040	-0.175	0.117	0.293*	0.050
Vocational education, male	-0.485*	0.076	-0.271*	0.033	-0.494*	0.086	-0.360*	0.042
Short cycle further edu., male	0.105	0.129	-0.241*	0.065	0.211	0.141	-0.322*	0.081
Medium cycle further edu., male	-0.011	0.122	-0.414*	0.065	0.058	0.135	-0.506*	0.080
Long cycle further edu., male	1.081*	0.101	0.011	0.068	1.358*	0.141	0.120	0.086
Couple has same level of edu.	-0.187*	0.074	0.222*	0.045	-0.172*	0.080	0.137*	0.040
Male more educated	0.122	0.099	0.108*	0.033	0.145	0.108	0.271*	0.055
Female income	0.269*	0.102	-0.166*	0.043	0.338*	0.110	-0.166*	0.050
Male income	-0.126*	0.065	-0.242*	0.028	-0057	0.068	-0.268*	0.032
Sickness, female	-0.060	0.109	-0.018	0.035	-0.087	0.116	-0.039	0.038
Sickness, male	-0.623	0.107	0.040	0.036	-0.117	0.115	0.028	0.040
Unemployment degree, female	-0.935*	0.171	0.147*	0.049	-0.946*	0.182	0.176*	0.056
Unemployment degree, male	-0.243	0.162	0.433*	0.056	-0.180	0.173	0.471*	0.065
Mass points (v_m^2, v_r^2)	-1.000*	0.258	-4.523*	0.741	2.157*	0.526	-2.352*	0.082
$p_1(v_r^1, v_m^1)$	0.762*	0.088			0.223*	0.090		
$p_2(v_r^2, v_m^1)$	0.114	0.067			0.482*	0.106		
$p_3(v_r^1, v_m^2)$	0.047	0.089			0.157	0.086		
$p_4(v_r^2, v_m^2)$	0.077	0.076			0.138	0.109		
$\mathrm{Corr}(v_r, v_m)$	0.413	0.284			0.222*	0.086		

Table 6: Full set of Results for Partnerships formed in the countryside

The higher divorce rates are not caused by the fact that we mismeasure cohabitation in the city. If because of the larger student population, spurious cohabitations more frequently take place in the city, this could explain the higher divorce rate there. However, (i) if we repeat our estimations excluding the cohabiting population, we find similar results[6] and (ii) the couples who move together to the city are likely to have a real relationship and we also find higher divorce rates for them. Finally, the results presented in Table 3 and 4 still hold when we change the definition of the city versus the countryside or also consider Aarhus and Odense to be cities. Including less populated areas in the city definition lowers the effect on dissolution risk in model 1, where the moving decision is not modelled. Not surprisingly, this effect also vanishes once moving is modelled explicitly.

4 Concluding remarks

In this paper, we investigated whether living in a less populated area lowers the divorce rate for a sample of Danish couples. We find, using the timing-of-events model of Abbring and van den Berg (2003), that conditional on location of marriage, the divorce risks are higher in the city but that this is for a large part caused by the sorting of relatively stable relationships in the countryside. This conclusion is confirmed by using an exclusion restriction.

References

[1] Abbring, J. and G. van den Berg (2003). "The Non-Parametric Identification of Treatment Effects in Duration Models", *Econometrica, 71, 1491-1517.*

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

[3] Bennett, N.G, A.K. Blanc and D.E. Bloom (1988). "Commitment and the Modern Union: Assessing the Link Between Premarital Cohabitation and Subsequent Marital Stability", *American Sociological Review*, 53, 127-138.

[4] Björklund A. and M. Sundström, "Parental Separation and Children's Educational Attainment: A Siblings Approach", IZA Discussion Paper 643 (2002).

[6] Since most partnerships begin as cohabitation (around 80%) the sample is severely reduced when we focus on those formally married. The effect of leaving Copenhagen is almost the same as when we included the cohabiting couples. The standard errors are however a lot larger due to the lower sample size which makes the results not as stastically robust as the ones based on the entire sample.

[5] Böheim, R. and J.Ermisch (2001). "Partnership Dissolution in the UK - the Role of Economic Circumstances", *Oxford Bulletin of Economics and Statistics*, 63, 2, 197-208.

[6] Burgess, E., H.J. Locke and M.M. Thomes (1963). "The family from institution to companionship, New York: *American books.*

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

[8] Calhoun, A.W.(1945). "A social history of the American family", Vol. 3. New York, *Barnes and Noble.*

[9] Chiappori, P.A., M. Iyigen and Y. Weiss (2005). "Marriage Markets, Divorce and Intra-household Allocations", mimeo, Columbia University.

[10] Christoffersen, H. (2003), Det danske bymønster og landdistrikterne, AKF, (In Danish)).

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

[12] 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.

[13] Drewianka, S. (2005). "A Generalized Model of Commitment", *Mathematical Social Science,* 52(3), 233-251.

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

[15] Forste, R. (2002). "Prelude to Marriage or Alternative to Marriage? A Social Demographic Look at Cohabitation in the U.S.", *Journal of Law and Family Studies*, 4(1), 91-104.

[16] Gautier, P.A., M. Svarer and Teulings C.N. (2005). "Marriage and the City", Working Paper 2005-01, Department of Economics, University of Aarhus.

[17] Gemici A. (2006), Family Migration and Labor Market Outcomes, Mimeo, NYU, 2006.

[18] Gruber J. (2000), Is Making Divorce Easier Bad For Children? The Long Run Implications of Unilateral Divorce", NBER Working Paper 7968 (2000)

[19] Jalovaara, M. (2001). "Socio-Economic Status and Divorce in First Marriages in Finland 1991-93", *Population Studies*, 55, 119-133.

[20] Kiefer, N.M. (1990). "Econometric Methods for Grouped Duration Data". In: Panel Data and Labour Market Studies, North-Holland.

[21] Lillard, L.A. (1993). "Simultaneous Equations for Hazards – Marriage Duration and Fertility Timing", *Journal of Econometrics*, 56, 189-217.

[22] Masters, A. (2006). "Marriage, Commitment and Divorce in a Matching Model with Differential Aging", Forthcoming in *Review of Economic Dynamics*.

[23] Peters, E. H. (1986). "Marriage and Divorce: Informational Constraints and Private Contracting", *American Economic Review*, 76(3), 437-454.

[24] Piketty, T. (2003), The impact of divorce on school performance: evidence from France: 1968-2002., CEPR discussion paper, 4146.

[25] Schultz, M.(1984). "Divorce in early America: Origins and patterns in three North-Central states", *The Sociological Quarterly*, 25 (Autumn 1984), 511-526.

[26] Statistics Denmark (2005), Vielser og Skilsmisser - børn i skilsmisser, (in Danish)).

[27] Sullivan, T. (1995). "Ex Ante Divorce Probability and Investment in Marital-Specific Assets: An Application to Home Ownership", Working Paper 95-5, University of Maryland, Department of Economics.

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

[29] Svarer, M. and M. Verner (2008). "Do Children Stabilize Relationships in Denmark?", *Journal of Population Economics,* 21, 395-417.

[30] van den Berg, G. (2001). "Duration Models: Specification, Identification, and Multiple Durations", in J.J. Heckman and E. Leamer, eds., *Handbook of Econometrics*, Vol. V, North Holland, Amsterdam.

[31] Weiss, Y. and R.J. Willis (1997). "Match Quality, New Information, and Marital Dissolution", *Journal of Labor Economics*, 15(1), S293-S329.

Appendix

A Likelihood function

Since we only observe the transitions on a yearly basis, we specify a model for grouped duration data (see e.g. Kiefer (1990)). The duration T_e, $e = r, m$ is observed to lie in one of K_e intervals, with the k_e'th interval being $(t_{k-1,e}; t_{k,e}]$ and the convention $t_0 = 0$ for $k_e = 1, ..., 15$. The probability that the duration T_e for an individual with explanatory variables $x_{e,t}$ and unobserved characteristics v_e is greater than $t_{k,e}$ given that the duration is greater than $t_{k-1,e}$ is given by:

$$P(T_e > t_{k,e}|T_e > t_{k-1,e}, x_{k,e}, v_e) \quad = \exp\left[-\int_{t_{k-1,e}}^{t_{k,e}} h_e(t|x_{e,t}, v_e)dt\right] \tag{2}$$

where $\Lambda_{e,k_e} = \int_{t_{k-1,e}}^{t_{k,e}} \lambda_e(t)dt$. The interval-specific survivor expression (2) is henceforth denoted by α_{e,k_e}. The probability of observing a given event in interval k_e, conditional on survival until $T_e > t_{k-1,e}$, is consequently $1 - \alpha_{e,k_e}$. If we do not specify a functional form for the baseline hazard within the interval, the $\Lambda_{k,e}$s are just parameters to be estimated.

Imposing the mixed proportional hazard formulation (1) and assuming that the observed covariates are time-invariant within intervals (i.e. years) – which implies that we only have to integrate over the baseline hazard – we can now express the interval-specific survival probabilities as

$$\alpha_{r,k_r} = \exp\left[-\exp\left[\beta_r' x_{r,k_r} + \delta D(t_m) + v_r\right] \cdot \Lambda_{r,k_r}\right]$$

and

$$\alpha_{m,k_m} = \exp\left[-\exp\left[\beta_m' x_{m,k_m} + v_m\right] \cdot \Lambda_{m,k_m}\right].$$

Notice, that $\Lambda = \int_{t_{k-1}}^{t_k} \exp(\lambda_i(t))dt$ is simply estimated as the average baseline hazard in the given interval. This corresponds to estimating a piecewise constant baseline hazard for each interval.

First, notice that each relationship contributes to the likelihood function as long as the relationship is intact. The contribution to the likelihood function from the relationship duration alone is therefore

$$\mathcal{L}_r = (1 - \alpha_{r,k_r})^{j_r} \alpha_{r,k_r}^{1-j_r} \prod_{l_r=1}^{k_r-1} \alpha_{r,lr},$$

where $j_r = 1$ if the relationship is not right-censored and 0 otherwise. Uncompleted durations therefore only contribute to the survivor probabilities. The interval indicator here runs monotonically from 1 up to the end of the relationship or is right-censored at k_r. The contribution for a given relationship is then $(1 - \alpha_{m,k_m})$ in intervals with a move and α_{b,k_b} in intervals without moves. Let the indicator variable, j_m, take the value 1 if a move occurs in a given interval and 0 otherwise. The contribution to the likelihood function from a move alone is then

$$\mathcal{L}_m = \prod_{l_m=1}^{k_r} (1 - \alpha_{m,k_m})^{j_m} (\alpha_{m,l_m})^{1-j_m}.$$

Combining the two expressions yields the full likelihood function

$$\mathcal{L} = \int \int \mathcal{L}_r \mathcal{L}_m dG(v_r, v_m),$$

where $G(v_r, v_m)$ is the joint distribution of the unobserved heterogeneity components. We use a flexible and widely applied specification of the distribution of the unobservables; it is assumed that v and v_m each can take two values, where one of the support points in each destination specific hazard is normalized to zero (i.e., $v_{r1} = 0$ and $v_{m1} = 0$), because the baseline hazard acts as a constant term in the hazard rates. Thus, there are four possible combinations of this bivariate unobserved heterogeneity distribution, each with an associated probability[7]. For more details on this class of mixture distributions in duration models, see e.g., van den Berg (2001).

[7]The four probabilities are: $P_1(v_r = 0, v_m = 0)$, $P_2(v_r = v_{r2}, v_m = 0)$, $P_3(v_r = 0, v_m = v_{m2})$, and $P_4(v_r = v_{r2}, v_m = v_{m2})$.