Location choices of highly educated foreign workers: the importance of urban amenities

Jan Rouwendal and Or Levkovich
Department of Spatial Economics, VU University, Amsterdam
This version: March 11th 2014

Abstract

Current developed economies' growth becomes increasingly dependent on the performance of innovation and skill-intensive industries. Therefore, the ability of cities to attract skilled or highly-educated individuals becomes more and more important for their growth and economic development. In this research we estimate a residential sorting model in order to shed light on the factors that determine the location choices of foreign skilled workers. We do so by estimating their valuation of various urban amenities in the municipalities of the densely populated Randstad area in the west of the Netherlands, and investigating which amenities increase the attractiveness of these municipalities. We also consider heterogeneity in individual preferences, and compare housing preferences and marginal willingness to pay for amenities between groups based on skill levels and origin. We find that job opportunities, accessibility, natural amenities and presence of historic monuments and buildings are highly valued by both domestic and migrants with high-education. Our results also provide evidence that social amenities, such as an existing community of migrants in a municipality, have an important role in determining the attractiveness of a location.

1 Introduction

Skills are generally thought to one of the drivers of economic growth. Human capital is an important determinant of productivity and 'the race between education and technology' (Goldin & Katz, 2009) is an important driver of the income distribution. Globalization has resulted in a decline of manufacturing in the western world and a shift of the core activities towards innovative and skill-intensive industries, which require human capital as one of its key resources. The increasing mobility, particularly of highly educated workers, underlines the need to make cities attractive places to work and live for such people. The ability of cities to attract international migrants, particularly those with higher human capital levels is therefore increasingly seen as an important indicator of their growth potential (Glaeser & Resseger, 2010; Glaeser & Saiz, 2003; Moretti, 2012).

The effect of immigration on economic growth has been studied extensively in recent literature, including a particular examination of the effects of different compositions of group of migrants, mostly based on skill and nationality diversification. Positive effects on productivity may result from the diversity in idea and skills, which may improve innovation (Hunt & Gauthier-Loiselle, 2008; Niebuhr, 2010; Ottaviano & Peri, 2006; Suedekum, Wolf, & Blien, 2009). Also, immigration flows of high human capital were also found to encourage local employment levels through a multiplier effect (Moretti & Thulin, 2013; Moretti, 2010). Moreover, presence of migrants may positively affect the utility of residing in a region by creating a multicultural environment which is perceived positively by a tolerant native population (Moretti, 2004; Niebuhr, 2010; Ottaviano & Peri, 2006; Suedekum et al., 2009). In contrast, negative impacts on productivity may rise from communication and cultural barriers, and a relatively high share of migrants group may also crowd-out natives in jobs, particularly those with medium skill levels (Eeckhout, Pinheiro, & Schmidheiny, 2010). Migration flows may also result in social tensions between groups, and may generate fear of foreign infiltration among natives (Bellini, Ottaviano, Pinelli, & Prarolo, 2008; Ottaviano & Peri, 2006; Ozgen, Nijkamp, & Poot, 2011; Suedekum et al., 2009).

Since the presence of migrants is so essential for urban and regional growth, it is important to know what makes a city an attractive place for highly skilled migrants. This research aims to shed light on this issue by considering the location choice of highly-educated foreign workers, and how their valuation of urban amenities differs from domestic workers. To do so, we apply a residential location-choice model to estimate the attractiveness of residential locations by low and high-skilled, domestic and foreign workers, and calculate and compare their willingness to pay for each of these amenities. Our location choice model is based on earlier work by Bayer et al (Bayer, Mcmillan, & Rueben, 2004), and is estimated on data over location choices of households in the Netherlands.

The paper is organized as following. In section 2 we will describe relevant and recent researches that studied the valuation of urban amenities and their role as attraction factors, and studies that investigated location choices of individuals, particularly skilled and migrant workers. In Section 3 we discuss the residential sorting model and explain the design of the estimation model. Section 4 describes the data that was used and the variables which were included. Section 5 describes the results of the estimation and section 6 provides discussion and concluding remarks.

2 The importance of urban amenities as an attraction factors for foreign skilled worker

2.1 Valuation of urban amenities

In recent years, increasing attention is given to the role of cities as centers of consumption as well as production. This perspective raises the importance of urban amenities and urban attraction factors to the growth of cities (Glaeser, Kolko, & Saiz, 2001). The perception of cities as consumption centers focuses the analysis of residential location choices on the provision of urban amenities, as they determine the unique characteristics of a city and the utility which individuals derive from them.

Glaeser et al (2001) find that high-amenity cities grew faster than low amenity cities, and they advise that urban policy should aim to attract workers on the basis of quality of life as well as on the basis of higher wages. They identify several critical urban amenities, such as presence of a variety of service and non-transportable consumer goods, physical setting and architecture, good public services and low transportation costs. They also find that amenities which attract human capital tend to be more important to urban growth, as growth in human capital results in an increase in productivity. The importance of urban amenities as attractors of workers is further studied by Adamson et al. (2004), which applied the Roback model (Roback, 1982) and examined whether skilled workers are more attracted to urban productivity or to urban amenities by studying the effects of urban agglomeration scale on the return to education. Urban productivity gains imply externalities of knowledge transfers, and higher demand for skilled workers which increases wage gaps for skilled workers. In contrast, urban amenities have an opposite effect on wages as they lead to an increase in the supply of labor. Their results show that net returns to education decline with urban scale.

Much research is still being conducted around the question of what determines valuation patterns of urban amenities. Valuation of urban amenities differs between groups in the population, based on their specific characteristics such as origin, skill and income levels. These differences influence location decisions of individual households between these groups. The implication is that households are sorted into different cities, or into neighborhoods within a city, based on the provision of certain urban amenities and their individual willingness to pay for each amenity. Brueckner et al (1999) finds that individuals in different income groups are located based on the spatial pattern of amenities in cities. They find that city centers with strong amenities tend to attract higher income groups. However, where city centre amenities are poorer or scarce, the rich prefer to locate in the city suburbs.

Van Duijn and Rouwendal (2013) also investigate residential sorting based on attractiveness of urban amenities, as they examine households' willingness to pay for cultural heritage in Dutch cities, differentiating between households with different individual characteristics. They consider that although historical centers are exogenous, they may attract other amenities such as shops and restaurant, which additionally contribute to the attractiveness of the city. This endogenous process is dealt with by historic instrumental variables. The results show that existence of an historical inner city contributes to a city's success, both directly or indirectly through related endogenous amenities. Marlet and van Woerkens (Marlet & van Woerkens, 2005) reach similar conclusions in their examination of the urban factors which attract the "creative class" in Dutch cities, as measured by growth in shares of employment in particular professions which were defined as "creative", following Florida (Florida, 2003). They also find that job opportunities as well as city-aesthetics, historic buildings and natural amenities contribute to growth of the creative class in Dutch cities.

2.2 Determinants of migrants' and skilled workers' location choices

The assertions that urban amenities are not valued equally by heterogeneous households, and that urban amenities can develop endogenously for social and economic conditions, were reinforced by the findings of the researches mentioned above. In the study of the location choices of migrants, these key issues are carefully addressed, and a greater emphasis is put on valuation differences between population groups and on the central role of social interaction as location amenity.

As in Van Duijn and Rouwendal and Marlet and van Woerkens, Rodriguez-Pose and Ketterer (2012) find that historical amenities, as well as nature amenities, are important in determining the attractiveness of regions. Particularly examining the role of urban amenities in the process of location choices of migrants in Europe, they also find that factors such as a presence of a large migrant community, regional wealth and favorable local labor market conditions are relevant in determining the geographical appeal for migrants of EU regions.

Aslund (2005), Damm (2012) and Jaeger (2006) study the location choices of international migrants, but they do not base the set of alternatives directly on urban amenities but on the effects of social interaction – the existing population demographics, and of the local labor market. Their findings emphasize the importance of existing migrants' community and network, and the expected labor outcome while choosing a migration destination. A local migrants' network is viewed as an attractive amenity for migrants, since it provides information about the local labor market at the destination and it assists new migrants in finding jobs (Åslund, 2005; Bauer, Epstein, Gang, & Al, 2007; Borjas, 1994; Jaeger, 2006; Munshi, 2003).

However, an existing community of migrants should also be considered as an endogenously determined amenity. Namely, a concentration of migrants in a city or neighborhood may be the result, as well as the cause, for migrants' location decisions. This was previously dealt with instrumental variable, as was done by Ottaviano and Peri (2006) and Mocetti and Porello (2010), which used the distance to immigration gateway as an instrument for current concentration of migrants. Damm (2012) addresses the endogenous concentration of migrants by exploiting a natural experiment in which refugees in Denmark were assigned quasi randomly to municipalities. Hunt and Gauthier-Loiselle (2008) and Niebuhr (2010) instrumented the existing share of skilled migrants by the lagged share of low-skilled migrants, as the shares of both groups are likely to be correlated (some urban amenities appeal to both), but is unlikely to be correlated with skilled migrants' effects on innovation and productivity.

Migrants' location decisions are also largely determined by their individual skill levels. Bartel (1989) focuses particularly on the subpopulation of highly-educated migrants and finds that their location decisions may be opposite to those of the rest of the migrants population. For example, she finds that highly skilled migrants from Europe or Asia to the US tend to relocate away from areas with a large community from the same ethnic origin. This finding shows that neglecting heterogeneity in personal characteristics within the group of migrants, particularly in skill levels, may lead to bias estimates of their choice of residential locations and their valuation of urban amenities.

Comparing location choices of migrant workers of different skill levels, Gottlieb et al (2006) find that doctorate degree holders have much higher valuation of regional amenities, even in comparison with other highly-skilled migrants. The authors explain this by arguing that doctorate holders have more bargaining power in employment negotiations, permitting them to demand and secure high amenities. In addition, they also find that educated migrants tend to value areas with higher percentages of university graduates, or better-educated cities. Contrary to Gottlieb et al.'s findings, Brown and Scott (2012) find that higher-educated workers (measured as academic degree holders) value recreational and destination labor market amenities similarly to the rest of the migrant population. They also find that compared with less-skilled workers, high skilled migrants place a higher value on the benefits of thick labor markets. This is explained as highly-educated workers are likely to seek larger labor markets in which they can specialize in their industry and occupation, increase their productivity and wages, and enjoy further accumulation of human capital from other specialized industry peers.

Despite the extensive research which was conducted in the field of valuation of urban amenities, and their attraction effect over skilled migrants, it is evident that the complexity of the issue leaves much room for further research. Researchers generally agree that residential decisions are closely dependent on several urban amenities, such as a large labor market, an existing community of migrants, urban scale, accessibility, natural aspects and historical amenities. However, careful attention must be given while addressing the urban amenities which are endogenously determined in the model, or that are correlated with unobserved variables. Furthermore, most studies do not provide estimations of the monetary value of the willingness to pay (either positive or negative) for amenities. The valuation of the marginal willingness to pay for urban amenities has a direct use in urban policy, but it may also help to make a comparison of the preferences and location decisions of groups of domestic and foreign, skilled and low-skilled workers. Understanding whether the decisions of highly educated migrants show more resemblance to those of the low-skilled migrants, or to those of the skilled native workers, would further help focus urban or regional policy that aims to attract skilled migrants.

3 The Residential sorting model

3.1 Methodology

The most prevalent method to estimate location choices in literature is by using the multinomial-logit (MNL) model. Although the model suffers from the independence of irrelevance alternative (IIA) property, allowing for a sufficient amount of heterogeneity among the consumers allows aggregate choice probabilities and substitution elasticities to be determined by the data.[1] The sorting model is based on the MNL model and is used to estimate the probabilities for each household *i* to choose each location *n*, based on the assumption of households' utility maximization. The sorting model is not only useful in estimating location choices of individuals, it also reveals the valuation of the alternative locations' characteristics, like hedonic regression models which are most commonly used in literature.[2] In additional to the average marginal willingness to pay for amenities, which is also revealed by hedonic analysis, the residential sorting model also provides the marginal willingness to pay for each observed population group.

[1] This was shown by McFadden and Train (McFadden & Train, 2000) for the mixed (or random coefficient) logit models, in which unobserved heterogeneity is included, but similar arguments can be put forward for a multinomial logit model that refers to a heterogeneous group of decision makers. See Bayer et al. (2004). See Gottlieb et al. (Gottlieb & Joseph, 2006) for an application of the mixed logit model to migration choices.

[2] See Bayer et al. (Bayer, Ferreira, & Mcmillan, 2007) for a comparison of the sorting model and hedonic price analysis.

Since house prices equalize demand and supply, they are likely correlated with the unobserved characteristics that affect location choices. Neglecting to consider these unobserved factors may results in bias in the estimation of choice probabilities and biased coefficients. Here, we follow Bayer et al.'s approach for instrument construction (Bayer et al., 2004) and construct the counterfactual share of migrants in a city, assuming no correlation exists with unobserved location heterogeneity, as will be explained further in later sections. The issue is most commonly dealt by constructing a price instrument which rises naturally from the model and is uncorrelated with unobserved neighborhood characteristics (Bayer et al., 2004; Klaiber & Phaneuf, 2010; van Duijn & Rouwendal, 2013). A similar phenomenon occurs in relation to endogenous amenities like the share of migrants located in a city. This share is also determined in part by the unobserved amenities, which results in an endogeneity problem. As mentioned, the concentration of migrants was previously instrumented by exogenous variables such as an historical immigrant "gateway", through which historical immigration flows have entered the country (Mocetti & Porello, 2010; Ottaviano & Peri, 2006). Although this variable might be suitable to explain location choices over a large geographical area like the United States, it is practically inapplicable in a small country like the Netherlands. Therefore, we construct an instrument for the share of migrants in municipalities by using the same procedure as in the construction of the price instrument. Namely, we construct the counterfactual share of migrants, as it would have been had there were no unobserved municipal characteristics which would have been correlated with the share of migrants. This is a valid instrument as it is highly correlated with the actual share of migrants, but by definition it is uncorrelated with the model's error term.

3.2 The residential sorting model – Design of the model

We consider a population of households i=1..*I*, that chooses a residential location out of a given set of alternatives n=1…*N*. Household i choose an alternative location n such that it maximizes its indirect utility, based on the provision of urban amenities *k=1..K* in each location.

$$MAX\ U_{i,n} = V_{i,n} + \varepsilon_{i,n} = \sum_{k=1}^{K} \alpha_{i,k} X_{k,n} + \varepsilon_{i,n} \quad , \quad U_{i,n} \geq U_{i,m}\ \forall\ n \neq m \quad (1.1)$$

Where $V_{i,n}$ denotes the indirect utility of household i from alternative n, $X_{k,n}$ denotes the value of the *k*-th characteristic of alternative *n*. $X_{k,n}$ includes all observed location characteristics, among them are also house prices and share of migrants in a location. $\alpha_{i,k}$ is a household type-specific coefficient which depends on households' individual characteristics:

$$\alpha_{i,k} = \beta_{0,k} + \sum_{l=1}^{L} \beta_{k,l}\left(Z_{i,l} - \overline{Z}_l\right) \quad (1.2)$$

Where $Z_{i,l}$ denotes the value of the *l*-th characteristic of household i, $\overline{Z}_l$ denotes the sample mean of characteristic l. Equation (1.1) can therefore be rewritten as:

$$U_{i,n} = \sum_{k=1}^{K} \beta_{0,k} X_{k,n} + \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \beta_{k,l} \left(Z_{i,l} - \overline{Z}_l \right) \right) X_{k,n} + \varepsilon_{i,n} \qquad (1.3)$$

The first expression on the right can be interpreted as the indirect utility of the "average household" from location *n*. The second expression is interpreted as the deviation from the mean indirect utility of location n of household *i*, based on its observed household characteristics. Hence, $\beta_{k,l}$ captures the cross effects between household and characteristics and urban amenities.

As we mentioned above, in practice it is difficult to assume that all relevant location characteristics are observed. Since it is plausible that households' utility is affected by other unobserved characteristics, an additional constant term ξ_n is introduced in the model in order to capture these specific location characteristics.

$$U_{i,n} = \sum_{k=1}^{K} \beta_{0,k} X_{k,n} + \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \beta_{k,l} \left(Z_{i,l} - \overline{Z}_l \right) \right) X_{k,n} + \xi_n + \varepsilon_{i,n} \qquad (1.4)$$

However, introducing the term ξ_n creates an additional problem, since the unobserved location characteristics are most likely to be correlated with the observed characteristics, particularly with housing prices. Neglecting to consider this would result a bias in the estimates of $\beta_{0,k}$. To address this endogeneity problem, we follow Berry et al's (Berry, Levinsohn, Pakes, & Berry, 1995) method and estimate the model in two steps. We begin by rewriting the model in equation (1.4) as:

$$U_{i,n} = \delta_n + \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \beta_{k,l} \left(Z_{i,l} - \overline{Z}_l \right) \right) X_{k,n} + \varepsilon_{i,n} \qquad (1.5)$$

With:

$$\delta_n = \sum_{k=1}^{K} \beta_{0,k} X_{k,n} + \xi_n \qquad (1.6)$$

In the first step of the estimation, we estimate equation (1.5) as a multinomial-logit model, assuming that the individual error is randomly drawn and independently and identically distributed (IID). The estimation is conducted by maximum-likelihood procedure, in which we estimate the vector of individual coefficients $\beta_{k,l}$ and the vector of mean indirect utilities from each location δ_n. In the second step of the estimation, we analyze the components of δ_n, by estimating equation (1.6) in a 2SLS model with instrumental variables.

The estimation of the first step is largely dependent on the equilibrium condition, according which the demand should equal the supply of houses in each location *n*. After defining the indirect utility function in (1.5), we calculate the probabilities of each household *i* to choose location *n* (denoted as $Pr_{i,n}$), by estimating it as a MNL model using a maximum-likelihood procedure:

$$Pr_{i,n} = \frac{e^{V_{i,n}}}{\sum_{n=1}^{N} e^{V_{i,n}}} \qquad (1.7)$$

The estimation of the first step results in a set of choice probabilities. Imposing an equilibrium restriction, we require that the sum of these choice probabilities would be equal to the existing housing stock in each location (S_n).

$$\sum_{i=1}^{I} Pr_{i,n} = S_n \tag{1.8}$$

The estimated coefficients $\beta_{k,l}$, which indicate the valuation of household with characteristic *l* for location characteristic *k*, and δ_n, which indicates the mean indirect utility from location *n*, are iteratively adjusted to reflect this equilibrium condition (Berndt, Hall, Hall, & Hausman, 1974).

In the second step, the estimated δ_n are further analyzed and are now explained by a set of location characteristics in a 2SLS regression. In this step we tackle the endogeneity issue which was mentioned before by instrumenting the price and share of migrants' variables. The computation of these instrumental variables will be discussed in the following section.

3.3 Endogeneity and the use of instrumental variables

In addition to the house prices variable, the share of migrants is also assumed to be correlated with the unobserved location characteristics term ξ_n. Therefore, the second step of the estimation includes instrumental variables for both. As was done by Bayer et al (2004) and Van Duijn and Rouwendal (2013), we construct both instruments based on the sorting model and existing data. We do so by assuming no unobserved neighborhood characteristics ($\xi_n = 0$). Following this, we simultaneously compute the price and share of migrants' vectors which would clear the market under this restriction. Intuitively, the price instrument is the set of prices that would prevail and result in equilibrium if the only amenities relevant for location decisions are observed in the model. Similarly, the computed share of migrants is such that would have been predicted if there were no unobserved attractive or unattractive amenities which affect the formation of a concentration of migrants. The computed instruments are valid since they are correlated with the original price and share of migrants' variables respectively, and they are defined such that they would have zero correlation with unobserved characteristics.

The instruments are constructed by first defining ($\xi_n = 0$), in equation (1.7), and then calculating the set of choice probabilities $Pr_{i,n}$ which arise from the revised equation:

$$\widehat{Pr_{i,n}} = \frac{e^{\sum_{k=1}^{K} \beta_{0,k} X_{k,n} + \sum_{k=1}^{K}\left(\sum_{l=1}^{L} \beta_{k,l}\left(z_{i,l} - \overline{z_l}\right)\right) X_{k,n}}}{\sum_{n=1}^{N} e^{\sum_{k=1}^{K} \beta_{0,k} X_{k,n} + \sum_{k=1}^{K}\left(\sum_{l=1}^{L} \beta_{k,l}\left(z_{i,l} - \overline{z_l}\right)\right) X_{k,n}}} \quad , \qquad \xi_n = 0 \tag{2.1}$$

These new set of probabilities must follow the equilibrium restriction, and therefore is plugged into equation (1.8), where we demand:

$$\sum_{i=1}^{I} ln\, \widehat{Pr_{i,n}} - ln\, S_n = \theta_n, \qquad \theta_n \cong 0 \tag{2.2}$$

The term θ_n refers to the difference between the computed demand and the actual supply of houses in location *n*, and is used to adjust the prices and share of migrants. The constants α_1, α_2 are used to moderate the iteration process.

$$\widehat{Prices_n^{it+1}} = Prices_n^{it} + \alpha_1 \theta_n \tag{2.3}$$

$$\widehat{Migrants_n^{it+1}} = Migrants_n^{it} + \alpha_2 \theta_n \tag{2.4}$$

In an iterative process, (2.3) and (2.4) are plugged back in equation (2.1) and the choice probabilities are calculated again. This process repeats itself until equilibrium is restored. After completion, the resulted prices and share of migrants vectors are then used to instrument for the respective variables in the second step of the estimation.

3.4 Spatial extensions

The spatial structure of the Randstad study area is of an urban cluster, which is characterized by high population density and continuity in the urban landscape. In many cities or towns, the municipal border is hardly noticeable, which raises the issue of spatial interdependence between the different municipalities. This issue has significant impact on location decisions, since households can reside in one municipality but still be able to enjoy the benefits of the neighboring municipalities, without experiencing high travel costs. To address this challenge we include a spatial lag in the explanatory variables, by adding a spatially-weighted average of amenity levels in the neighboring municipalities ($PX_{i,k}$). The weights are determined by a row-standardized inverse-distance between municipality *i* to all other municipalities $j=1...n,\ j\neq i$.

$$PX_{i,k} = \sum_{j=1, j\neq i}^{n} \frac{\frac{1}{d_{i,j}}}{\sum_j \frac{1}{d_{i,j}}} * X_{j,k}$$

Additionally, we also consider the fact the unobserved amenities may also be spatially correlated, and therefore, following Anselin (Anselin, 1988) and Anselin et al. (Anselin, Bera, Florax, & Yoon, 1996) we test for spatial correlation in the residuals of the model using Moran's I and Lagrange-multiplier test (see results in table 1).

Table 1 - Test statistics for spatial dependence

	Statistic	p-value
Moran's I	-0.051	0.419
Lagrange multiplier	0.033	0.855
Lagrange multiplier (Robust)	0.076	0.783

The statistics values obtained for Moran's I and the Lagrange-multiplier test for residual spatial correlation show little evidence that such correlation exists. Although these values suggest that we can estimate the model without considering spatial correlation in the model's residuals, we include an additional estimation of the model using Drukker et al.'s (Drukker, Egger, & Prucha, 2013; Drukker & Prucha, 2011) GMM/IV estimation method, for the purpose of robustness check. The GMM/IV is a two step estimation of spatial autoregressive disturbances model with consideration in endogenous regressors, and its formulation can be found in Drukker et al. (2011, 2013).

3.5 Calculations of the marginal willingness to pay for urban attributes

The estimation of the sorting model produces a set of coefficients that determine the mean valuation and household specific valuation of each of the observed attributes. We use these estimated coefficients to compute the average and marginal willingness to pay for the observed location attributes. The marginal willingness to pay for characteristic *n* by a household belonging to group *l* is the change in the price that keeps utility constant after a small change in the value of the *k*-th characteristic:

$$\frac{\delta P_n}{\delta X_{k,n}} = \frac{\left(\beta_{0,k}+\sum_{l=1}^{L}\beta_{k,l}\left(Z_{i,l}-\overline{Z_l}\right)\right)}{\left(\beta_{0,p}+\beta_{p,l}\left(Z_{i,l}-\overline{Z_l}\right)\right)} P_n \tag{3.1}$$

Household specific preferences are designed to have zero mean, this implies that the MWTP of the average household is:

$$\frac{\delta P_n}{\delta X_{k,n}} = \frac{\beta_{0,k}}{\beta_{0,p}} P_n \tag{3.2}$$

Computing the average and household type-specific MWTP, we are able to estimate the monetary value which is placed on various urban amenities by each group in the population, based on skill level and origin. This allows measuring which amenities are perceived to be more valuable and attractive to each of the groups, explaining their location patterns and comparing the differences in preferences for urban amenities between them.

4 Data and study areas

4.1 Databases

For residential location choices data we use Netherlands housing research survey (WoON) 2012, which was conducted as a joint co-operation between the Ministry of the Interior and Kingdom Relations (BZK) and the Dutch Central Bureau of Statistics (CBS). Data over municipality characteristics is also taken from the Dutch CBS.

4.2 Randstad municipalities analysis

The western area of the Netherlands is characterized by high levels of urbanization, and it is in fact an agglomeration of cities (the *"Randstad"*). Among the cities which are included in the Randstad are the four largest cities in the Netherlands – Amsterdam, Rotterdam, Den-Haag (The Hague, also known as *'s-Gravenhage*) and Utrecht. The region also includes other populated municipalities such as Almere, Zaanstad, Amersfoort, Leiden, Zoetermeer and Dordrecht. Despite its high population, the centre of the Randstad remains relatively rural (*"De Groene Hart" – The Green Heart*), and its municipalities maintain an agricultural character. The borders of the Randstad are not officially specified, and therefore in this analysis we included 135 different municipalities (see appendix A for the full list), all within a short commuting distance from the main population centers in the four largest cities (see map in figure 1).

The fact that the Randstad is relatively small and urbanized implies that commuting is common, and that individuals may live in one municipality, but work and enjoy amenities in other nearby municipality. To address the issue of spatial interdependence we introduce spatial data in the model using spatial matrixes, which were constructed based on contiguity of neighboring municipality as well as on inverse distance between municipality centroids.

4.3 Variables included – Household characteristics

Households are identified based on several characteristics, as indicated by each respondent in the WoON 2012 survey. We identify respondent's age, a dummy variable indicating whether the household includes children, and the household's income in logarithm. We also add two additional dummy variables indicating whether the respondent is a migrant, and whether she is skilled. The identification of a respondent's migrant status or skill level is somewhat uncertain. While skill level is more clearly determined, as we define a person as "skilled" whether he or she indicates to have obtained at least a professional or university degree, inaccuracies may still exist as some may be more inclined to indicate a different skill level. For instance, students in advanced progress of obtaining a degree may indicate themselves as degree holders.

This issue becomes more complicated with the case of migrants status. We identify respondents as 'migrants' if they have indicated themselves to be not of Dutch descent. This broad definition includes recent immigrants, but it also includes those who were born in the Netherlands and whose parents (or perhaps grandparents) arrived to the Netherlands as immigrants. The definition of a respondent as a native may also be inaccurate. For example, respondents who were not born in the Netherlands but were living there for many years may view themselves as natives. This issue leaves much area for interpretation by respondents and researchers alike, and makes the identification of the group of migrants quite difficult. Nonetheless, we may argue that this definition can be viewed as an advantage, and it may even better serve the purpose of the research. The reason for that is that when a respondent views herself as a migrant regardless of whether she was born in the Netherlands or not, her preferences in housing may be more similar to those who broadly define themselves as migrants rather than those who define themselves as natives. In this sense, the self-identification as migrants may prove to be more suitable for the research.

Figure 1: The Randstad study area within the Netherlands.

According to WoON 2012 survey data, approximately 24% of the survey respondents who live in the Randstad area are of foreign origin (6,322 respondents). Areas with the highest concentration of migrants are Amsterdam metropolitan area (*Groot Amsterdam)* and The Hague metropolitan area (*Agglomeratie 's-Gravenhage),* both with approximately 31% of respondents indicating themselves as migrants. Areas with the lowest concentration of migrants were Delft-Westland and East Zuid-Holland *(Oost Zuid-Holland)*, with 11% and 13% migrants respectively.

The Randstad area also has a large proportion of skilled respondents; approximately 45% of the skilled respondents in WoON 2012 have reported to live in one of the Randstad municipalities.

Out of the skilled population in the Randstad, approximately 21% (or 1,832) respondents are skilled migrants. Skilled migrants also form 29% of total migrant respondents in the province.

4.4 Variables included – Urban amenities

In order to explain the location choice of different population groups in the Randstad area, we included several amenities as explanatory variables. Following the findings of previous researches, we included recreational, cultural and natural amenities, as well as amenities which stress labor possibilities, accessibility and urban scale. As a proxy for existence of migrant networks and presence of cultural and ethnic goods, we also include the share of migrants in each municipality as explanatory variable. Average municipality housing prices are also included in our analysis. Both prices and the share of migrants are instrumented as explained in section 3.3.

Housing prices are calculated using hedonic regressions, which were based on data gathered from NVM, the Dutch real estate association. Data over the share of migrants in each municipality is taken from the Dutch Central Bureau of Statistics (CBS). According to the Dutch CBS definition, a migrant is defined as a person which at least one of his/her parents was not born in the Netherlands. This broad definition is quite different from the WoON data definition of a foreigner, according which respondents are required to indicate themselves whether they consider themselves as belonging to a non-Dutch ethnic group. Although the differences between definitions may create problems, they are not critically restrictive since the CBS definition is used to determine the share of a community of migrants as an alternative location's characteristics, while the WoON definition is used to characterize individuals' preferences.

Consistently with the discussion presented in the literature regarding the possible positive and negative effects of an existing foreign community, we introduce the share of migrants in the model both directly and as a square term. The purpose of this specification is to identify whether the marginal effect of an increase in the share of migrants differs between lower and higher levels of shares. In accordance with the literature, we expect that the positive effects of an existing community of migrants are dominant when the share of migrants is relatively low, and the negative become dominant where the share of migrants is higher.

To represent culture and recreation we use the number of official monuments in the municipalities (see appendix A), which have shown by Van Duijn and Rouwendal (2013) to be a good proxy for historical and cultural heritage, as well as other recreational and commercial activities which are attracted to historical scenery like cafes and restaurants. The data is taken from the Dutch Cultural Heritage Agency (RCE). We also use the area of nature coverage in a municipality as an indication for natural scenery amenity in a municipality, which is assumed to have positive valuation due its recreational and aesthetic attributes.

Since previous researchers found that labor market conditions matter for location choices of migrants, particularly with high education, we added the number of jobs in a municipality, and its level of accessibility (see appendix A). We also included a rail-accessibility variable, which is measured by the average distance of all residents in an area to the nearest train station (available from the Dutch CBS). Much like the number of jobs variables, which represents the “thickness” of the labor market, we also add the location-quotient of employment in information and communication (ICT) industries. A relatively high concentration of these knowledge intensive industries (approximately 60% of employees in this sector have a university or professional degree) in a specific municipality is assumed to be particularly appealing for skilled workers of foreign origin. The ICT sector is over represented (LQ of 3.47) in the Den-Haag metropolitan area, Het Gooi (North-eastern Randstad), Utrecht, Amsterdam and Haarlem areas. In the rest of the sub-regions of the Randstad, Notably in Groot-Rijnmond where Rotterdam is located, ICT industry is quite underrepresented (see appendix A).

Table 2 - Correlation between urban amenities variables

	ln(price)	Perc. Mig	Accessibility	Jobs	Monuments	nature	LQ (ICT)	Number of Households
ln(price)	1.00							
Perc. Mig	0.06	1.00						
Accessibility	-0.31	-0.47	1.00					
Jobs	0.02	0.68	-0.19	1.00				
Monuments	0.19	0.45	-0.11	0.79	1.00			
nature	0.44	0.10	-0.10	-0.07	-0.05	1.00		
LQ (ICT)	0.42	0.26	-0.22	0.13	0.10	0.21	1.00	
Number of Households	0.00	0.70	-0.20	0.99	0.80	-0.06	0.12	1.00

In addition, we also attempt to proxy the urban scale of a municipality using the number of households variable. As was previously noted by Gottlieb et al. (2006), who include city size as a similar variable to reflect urban scale, this inclusion creates a problem as urban scale is correlated both with the provision of recreational amenities (such as theatres, restaurants and major sport leagues) and with a thick labor market. Therefore, its coefficient is likely to have a positive bias. Examining the variables’ correlation table (see table 2) we find that the number of households is too correlated with most variables to be included in the estimation. Moreover, due to its high correlation with the number of jobs (99%), both variables may serve as a proxy for urban scale. The fact that urban scale may be embodied in many variables further emphasizes the problem of endogeneity in urban amenities variables, and the concern that it is difficult to identify amenities which are independent from other urban attraction factors.

5 Results

5.1 First step estimation results

Table 3 describes the variables which were used in the estimation of the sorting model, as explained in sections 3 and 4.

Table 3 - Specification description

individual characteristics (Z_l)	Alternative characteristics (X_k)	Spatial lags in the explanatory variables (PX_k)
• Age of respondent • Dummy if has kids • Income • Skilled native dummy • Low-skilled migrant dummy • Skilled migrant Dummy	• ln (Price) • Share of migrants • Share of migrants square) Accessibility (Distance from intercity train station • Number of jobs (‘000) • Monuments • LQ (ICT) • Percentage of nature coverage.	Included for all alternative characteristics variables except for prices.

The results of first step include the coefficients of the cross effects between individual and location characteristics, as well as the vector of location-specific constants, which indicate the indirect utility of the mean household from each of the alternatives. The estimation shows that the coefficients of the households-amenities cross effects are different from each other (see table 4). This serves as a preliminary demonstration that valuation of urban amenities differs between different subgroups of the population.

Examining the results of first step estimation, we first observe that skilled respondents have a positive and significant cross effect with prices, compared with the other groups. This does not necessarily suggests that these groups value higher housing prices as a location attribute, but rather that these groups are less sensitive to housing prices. The results also show positive and significant cross coefficients of existing community of migrants among both skilled natives and migrant respondents of both skill levels, where the migrants’ cross coefficient is higher. This indicates that these groups are likely to have a higher than average valuation of strong presence of migrants in a municipality. Introducing the square term of migrant community share, we see a small negative and significant coefficient among migrants, indicating that higher shares of migrants in a municipality may result in a negative valuation.

Table 4 - First step results

	age	Kids dummy	ln (income)	Native Skill	Mig-Lowskill	Mig-Skill
ln_Price	**-0.001918** (0.004)	**-0.928847***** (0.1482)	**-0.523011***** (0.08986)	**2.064872***** (0.16167)	**-0.250853** (0.18146)	**1.532261***** (0.22564)
P.mig	**-0.0014***** (0.00015)	**-0.045818***** (0.00552)	**-0.022187***** (0.0032)	**0.059102***** (0.00603)	**0.10359***** (0.00747)	**0.119127***** (0.00987)
P.mig (square)	**0.00002***** (0)	**0.000619***** (0.00012)	**-0.000019***** (0.00007)	**-0.000492***** (0.00014)	**-0.00098***** (0.00016)	**-0.00128***** (0.00021)
Accessibility (dist. from station)	**-0.000666***** (0.00014)	**-0.01568***** (0.00492)	**0.003155***** (0.00345)	**-0.01007***** (0.00552)	**0.001466***** (0.00722)	**-0.057929***** (0.00955)
Jobs	**-0.000057***** (0.00001)	**-0.001145***** (0.00033)	**0.000081***** (0.00018)	**-0.000318***** (0.00034)	**0.000291***** (0.00038)	**-0.00184***** (0.00047)
Monuments	**0.000003***	**0.000059***	**0.000072***	**-0.00009***	**0.00006***	**0.000104***

	(0)	(0.00002)	(0.00001)	(0.00002)	(0.00002)	(0.00003)
Nature	**0.000182*** **	**-0.006342*** **	**0.006819*** **	**-0.007417*** **	**0.00479*** **	**-0.005422*** **
	(0.00003)	(0.00115)	(0.00068)	(0.00119)	(0.00145)	(0.00181)
LQ (ICT)	**0.002918*** **	**-0.028096** **	**0.095225** **	**-0.004716** **	**0.051952** **	**0.107249** **
	(0.00051)	(0.01968)	(0.01182)	(0.02239)	(0.0226)	(0.02888)
P.mig (Spatial lag)	**0.009091*** **	**0.32805**	**0.190527***	**0.405742**	**-0.183939**	**0.944389**
	(0.00305)	(0.11735)	(0.08035)	(0.13562)	(0.15011)	(0.23472)
P.mig (square) (Spatial lag)	**-0.000165*** **	**-0.007726*** **	**-0.003758*** **	**-0.011449*** **	**0.004802*** **	**-0.018607*** **
	(0.00008)	(0.00314)	(0.00211)	(0.00357)	(0.004)	(0.00598)
Accessibility (dist. from station) (Spatial lag)	**0.002432*** **	**0.033196** **	**0.054198** **	**0.115608** **	**-0.073399***	**0.335552***
	(0.00101)	(0.03779)	(0.02592)	(0.04215)	(0.0522)	(0.06766)
Jobs (Spatial lag)	**-0.000092*** **	**0.008362*** **	**-0.002954*** **	**0.03011*** **	**-0.01474** **	**0.002237** **
	(0.00023)	(0.00896)	(0.00569)	(0.00983)	(0.01104)	(0.01457)
monuments (Spatial lag)	**0.000089*** **	**0.001479*** **	**0.002324*** **	**-0.004813*** **	**0.003272*** **	**0.002602*** **
	(0.00001)	(0.00057)	(0.00036)	(0.00065)	(0.00069)	(0.00085)
nature (Spatial lag)	**0.000596*** **	**-0.0174*** **	**0.002992*** **	**-0.00208** **	**0.010198** **	**0.009521** **
	(0.00024)	(0.0094)	(0.00631)	(0.01026)	(0.01203)	(0.01634)
LQ (ICT industries) (Spatial lag)	**-0.02918*** **	**0.299724**	**-0.53669***	**0.600603**	**-0.807229**	**-0.224861**
	(0.00273)	(0.10679)	(0.06698)	(0.11847)	(0.13023)	(0.15456)

N=27163
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: values represent the coefficient of the corresponding interaction variables.

Examining the cross coefficients of the number of jobs in the municipality, values seem to be relatively low, particularly among skilled native who have a negative cross coefficient, indicating a lower sensitivity to the number of jobs or to the urban scale of the municipality. These results can be explained by the cross coefficients of the spatial lag, or the number of jobs in neighboring municipalities. Expectedly, the values of these coefficients are generally higher, especially for the skilled groups of both local and migrant respondents. This corresponds with higher preferences of skilled workers to reside in proximity to a large labor market and enjoying its urban-scale related amenities, without actually residing there.

Regarding the valuation of a concentration of ICT employment, the results show positive and significant coefficients among migrants of both skill level groups, with a much stronger valuation among skilled migrants. The results also show a negative and significant coefficient among skilled respondents and a positive coefficient among respondents with higher income levels. Since the ICT industry is skill-intensive, this result is somewhat surprising. One possible explanation is that the LQ variable refers to a COROP area (NUTS3), which incorporates several municipalities. This means the industry may be concentrated in only one municipality within the COROP area, which would make the rest of the municipalities in the COROP less attractive for skilled workers.

Cross coefficients of monuments show that all groups are relatively insensitive to these amenities. This somewhat changes when we examine the cross coefficients of the spatial lags, or the number of monuments in neighboring municipalities. The spatial-lag results show that among the migrants groups there is a stronger preference for monuments in surrounding municipalities rather than in the municipality of residence itself. These results are also repeated in the preference for natural and recreational amenities, where among all groups preferences for natural amenities in neighboring municipalities are found to be stronger.

5.2 Second step estimation results

Table 5 describes the results of three model specifications – The un-instrumented OLS model, the 2SLS model where prices and share of migrants are instrumented, and the GMM/IV model for spatial autocorrelation in the residual, in the presence of endogenous regressors.

As expected, the results of the OLS model (column 1) show lower statistical significance levels of the estimated coefficients, and their values quite differ from those of the other specification. Moreover, while examining the results of the GMM/IV specification, we see that as predicted by the spatial statistics indicators, the spatial correlation in the error term is indeed relatively weak (Spatial rho = 0.07). These two findings bring us to the decision to focus the analysis on the 2SLS model (column 2), which still considers spatial dependence in the explanatory variables.

The 2SLS model results which are reported in column 2 show several important findings. First, as expected, price and distance from train stations coefficients were found to be negative, while coefficients for number of jobs, monuments and nature coverage were found to be positive. This indicates that the corresponding amenities were found to be negative and positive respectively. As for the percentage of migrants, the straightforward term was found to have a positive effect, while the square term was found to have a small but significant negative coefficient. The interpretation of this finding is that the percentage of migrants is found to be valued positively by the average household, but its valuation decreases in higher levels.

Table 5 - Second step results

	(1)	(2)	(3)
variable	**OLS (se)**	**2SLS (se)**	**GMM/IV (se)**
ln_Price	-1.3412 (0.473553)***	-5.2204 (0.592238)***	-5.2055 (1.195986)***
P.mig	0.163 (0.02166)***	0.172 (0.027089)***	0.1711 (0.027197)***
P.mig (square)	-0.0032 (0.000575)***	-0.0039 (0.000719)***	-0.0039 (0.000748)***
Accessibility	-0.0086 (0.013818)	-0.0184 (0.017281)	-0.0185 (0.017646)
Jobs	0.0127 (0.001923)***	0.011 (0.002405)***	0.0109 (0.002445)***
Monuments	-0.0001 (0.000123)	0.0004 (0.000154)**	0.0004 (0.000195)*
Nature	0.0038 (0.003834)	0.0168 (0.004795)***	0.0167 (0.005925)***
LQ (ICT)	-0.0473 (0.07795)	0.0212 (0.097486)	0.0097 (0.099386)
P.mig (Spatial lag)	0.6215 (0.363302)*	1.3812 (0.454356)***	1.3378 (0.502419)***
P.mig (square) (Spatial lag)	-0.0149 (0.009951)	-0.0274 (0.012445)**	-0.0264 (0.013068)**
Accessibility (Spatial lag)	0.0705 (0.105852)	0.0853 (0.132382)	0.0832 (0.13471)

Jobs (Spatial lag)	0.0318 (0.026278)	0.0161 (0.032865)	0.0147 (0.033811)
monuments (Spatial lag)	-0.0025 (0.001699)	-0.0001 (0.002125)	-0.000025 (0.002288)
nature (Spatial lag)	-0.051 (0.035821)	-0.1143 (0.044798)**	-0.1111 (0.048484)**
LQ (ICT industries) (Spatial lag)	0.361 (0.424465)	0.7951 (0.530848)	0.8264 (0.554267)
Constant	6.0811 (5.795729)	45.8762 (7.248299)***	46.0308 (12.976206)***
Price level instrument	No	Yes	Yes
Share of migrants instrument	No	Yes	Yes
			Spatial Rho = 0.0538
n = 135 Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1			

The results of the spatially-lagged amenities are found to be statistically insignificant for most amenities. Among the spatial-lag variables, the percentage of migrants' variables stands out. It appears that these coefficients are both statistically significant, as well as estimated with much higher values compared with the non-lagged variables. The interpretation of this finding is that having a large percentage of migrants in nearby municipalities is valued very positively, and it increases the expected attractiveness of a municipality much more compared to having large percentage of migrants in the municipality of choice itself. The spatially-lagged variable of the percentage of migrants squared also has a higher coefficient value compared to the non-lagged variable. This suggests that also under spatial lags the positive effect of a large community of migrants is declining in higher levels.

5.3 Marginal willingness to pay for urban amenities

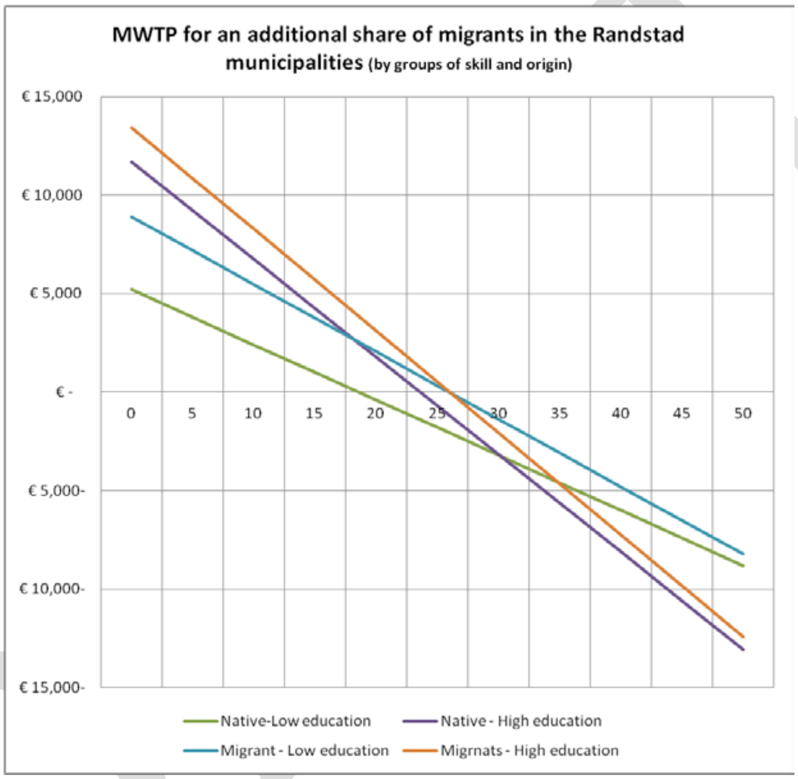
Having estimated the two steps of the residential sorting model, we are able to use the coefficients in order to calculate the marginal willingness to pay for each of the amenities included (as specified in section 3.4). The calculation is based on the coefficients obtained by the estimation of the 2SLS, as specified in column 2 of table 5. In the process of calculation we distinguish the valuation of each of the amenities by dividing the sample into four groups based on skill and origin- low-skilled natives, high-skilled natives, low-skilled migrants and high-skilled migrants.

Table 6 - Willingness-to-pay for a one percent increase of migrants in the municipal population (in Euro)

P.mig	AWTP	Native-Low education	Native - High education	Migrant - Low education	Migrants - High education
0	7,561.7	5,205.9	11,709.0	8,913.2	13,470.8
5	5,837.5	3,804.2	9,230.7	7,198.6	10,878.8
10	4,113.4	2,402.5	6,752.3	5,484.0	8,286.8
15	2,389.2	1,000.7	4,274.0	3,769.4	5,694.8

20	665.0	-401.0	1,795.6	2,054.8	3,102.8
25	-1,059.1	-1,802.8	-682.7	340.2	510.8
30	-2,783.3	-3,204.5	-3,161.1	-1,374.4	-2,081.3
35	-4,507.5	-4,606.2	-5,639.4	-3,088.9	-4,673.3
40	-6,231.6	-6,008.0	-8,117.8	-4,803.5	-7,265.3
45	-7,955.8	-7,409.7	-10,596.1	-6,518.1	-9,857.3
50	-9,680.0	-8,811.4	-13,074.5	-8,232.7	-12,449.3

Figure 2: MWTP for an additional share of migrants in Randstad municipalities (by groups of skill and origin).



Since we are not able to distinguish migrants' origin, we cannot determine whether decision patterns differ between migrants' from different countries of origins. Nevertheless, our finding partially contradict Bartel (1989), who found that skilled migrants of certain origins to the US tend to relocate away from large concentration of migrants, while low-skilled migrants tend to remain in areas where there is a concentration of migrants. We find evidence that suggests that the tendency of migrants to concentrate in certain locations depends on their skill level, as well as on the existing concentration of migrants in each location.

Moreover, the WTP for concentration of migrants is not necessarily always highest among groups of foreigners. For instance, where the share of migrants is low (below approximately

15%-20%) the willingness to pay for presence of migrants is higher among skilled natives compared with that of low skilled migrants.

Investigating the average and marginal willingness-to-pay for the rest of the examined urban attributes, we find several additional differences in valuation between the various groups (see table 7a).

Table 7a - Marginal Willingness-to-pay for urban amenities (in Euro)

	AWTP	Native-Low education	Native - High education	Migrant - Low education	Migrants - High education
Jobs	485.1	442.1	666.9	434.8	502.0
Monuments	15.7	14.4	16.8	16.0	25.1
Nature	735.8	719.1	661.5	870.7	686.1
Accessibility *	-808.3	-480.2	-1,363.4	-404.8	-3,763.1
LQ (ICT industries) *	924.5	252.8	104.1	2,210.8	6,103.7

** Based on coefficients which are not statistically significant*

While skilled domestic workers have a relatively high MWTP for jobs (667 EUR for an additional 1000 work places), the rest of the groups have a relatively similar MWTP values, around 430-500 Euros for an additional 1000 jobs. Although skilled migrants have a higher valuation for number of jobs (or urban scale), compared with both groups of low skilled workers, their MWTP is not sufficiently higher in order to conclude that there is a skill-bias in the valuation of this amenity. On the other hand, the valuation of natural and recreational amenities supports the assertion that a skill-bias exists in valuation pattern of urban amenities. Educated workers of both origin groups present similar MWTP for nature coverage (661 and 686 EUR for every additional square kilometer, for local and foreign workers respectively). This is while lower-educated workers show a much higher MWTP of 719 and 870 EUR for locals and migrants respectively.

The marginal willingness-to-pay for monuments or historic buildings presents a similar pattern. Consistently with the findings of Van Duijn and Rouwendal (2013), the highest values of MWTP are among highly-educated respondents. However, the gaps between the values are relatively large. The highest MWTP by a margin is among skilled migrants (25 EUR for every additional monument), while the second highest is 17 EUR among educated natives. Lower-educated native and migrants workers have a similar valuation of 14.4 and 16.0 respectively, values which are only slightly lower than that of the skilled natives' group. The results show that highly-educated households generally still have a higher MWTP for historic buildings compared with low-educated households, but mainly that skilled migrants have a clearly distinct preference for historic cities. Despite the fact that these WTP for monuments values appear relatively low at first sight, they are particularly high when reminded about the number of monuments in Dutch municipalities (described in appendix A). 21 municipalities in the sample have more than 200 monuments, and the municipalities of Utrecht, Leiden, Haarlem and Den-Haag all have over

1,000 monuments. This implies that the average household is willing to pay around 20,000 Euros to reside in one of these municipalities due to their provision of monuments and historical buildings alone. The WTP for monuments extremely increases when considering Amsterdam, which has 7,442 listed monuments – between 106 and 186 thousand Euros. Arguably, this high number seems particularly high compared to initial expectations. A possible explanation for this can be attributed to the indirect effects of cultural heritage. As mentioned before, monuments and historical centers are also correlated with other unobserved urban amenities, such as commercial and leisure consumption activities. These factors have an indirect contribution to the positive effect of monuments, and are reflected in the high willingness-to-pay values that are measured here. Given the number of monuments in the sample, the differences in WTP values between population groups are sharpened. For example, a skilled migrant is willing to pay approximately 36,000 Euro in order to reside in the city of Utrecht due to its 1,460 monuments. In contrast, a low skilled native is willing to pay 21,000 Euro for Utrecht's historical heritage, a difference of 15,000 Euro in valuation.

Average MWTP values of other labor related amenities, such as accessibility level and concentration of ICT industries, are harder to interpret due to the fact that they are based on statistically insignificant coefficients, and therefore the values are susceptible to bias. However, the statistically significant coefficients which were obtained in the first stage of the estimation still allow us to determine group-relative valuation of these amenities by examining deviations from the mean (see table 7b).

Table 7b - Marginal Willingness-to-pay for urban amenities (deviations from mean values in Euro)

	AWTP	Native-Low education	Native - High education	Migrant - Low education	Migrants - High education
Jobs	484.2	-43.0	181.8	-50.3	16.9
Monuments	15.7	-1.4	1.1	0.3	9.4
Nature	735.9	-16.8	-74.3	134.8	-49.7
Accessibility	-808.3*	328.2	-555.1	403.6	-2,954.8
LQ (ICT industries)	924.5*	-671.7	-820.4	1,286.3	5,179.2

** Based on coefficients which are not statistically significant*

The values in table 7b show preference towards accessibility is strongest among skilled workers of both origin groups. Although it is particularly strong among educated migrants as they have the highest negative valuation by a large margin for an additional average kilometer distance from train station. Moreover, highly educated migrants also have a stronger willingness to pay for residing where a concentration of knowledge-intensive ICT industries exists. This is particularly interesting when compared to WTP of other groups – contrary to the expectations, educated natives have below average valuation of this amenity, while low-educated migrants have a positive valuation. This discrepancy can be partially explained by the fact the Location quotient measures were only available at a COROP area level (NUTS3), which may have resulted in a bias in the estimated values.

6 Discussion and conclusion

The results of our estimation of the sorting model supports the findings of previous researches, as we find evidence that residential location choices in Dutch municipalities depend much on the provision of urban amenities. Moreover, our results also provide further evidence that valuation patterns of urban amenities differ between individuals based on their household characteristics, particularly origin and skill level. We find that that job opportunities, as well as accessibility, natural amenities and particularly historical city centers, play a significant role in raising the attraction power of municipalities. Additionally, the role of social amenities, such as the share of migrants in a municipality, is also found to be an important explanatory variable. The implication is that cultural diversity and social interaction between inhabitants have a strong impact on households' location decisions, and on the attractiveness of municipalities, at least as much as other consumer or labor-related urban amenities. This positive effect on municipality attractiveness is found to be even stronger in municipalities neighboring large communities of migrants, as reflected by the higher value of the spatially-lagged variable's coefficient. Consistent with findings of previous researches, our results also show that both positive and negative effects may result from an existent community of migrants. We observe this as the MWTP for an additional share of migrants in a municipality changes with the share of migrants in the municipalities- it is found to be high among all skill and origin groups where migrants shares are low, but decreases where the shares and presence of foreigners is relatively high.

We also find that with respect to most urban amenities, the location preferences of skilled migrants are generally more similar to those of the highly-educated natives rather than to low-skilled migrants, assuming that other household characteristics are identical. This is visible in MWTP values of both work-related amenities, such as the number of jobs or accessibility, but also in preferences for other urban amenities, such as historical buildings and nature. In this sense, and also somewhat expectedly, preference for an existing community of migrants forms an exception in the skill-based pattern of preference for urban amenities. Skilled migrants are found to be more similar to low skilled migrants in their preferences for shares of migrants in a municipality. Their MWTP values for an existing community of migrants even exceed those of the low-skilled migrants, assuming all other characteristics are identical. This may be explained as migrants of all skill levels are assumed to derive utility from cultural goods, but skilled migrants may also enjoy the network benefits of a large community of migrants, which may increase their expected labor outcomes. Furthermore, our results point out that highly-educated migrants have the highest MWTP among all groups for historic city centers and buildings. This valuation becomes considerably significant in municipalities that are rich in historical districts.

These findings may be useful for urban policy that aims to attract foreign educated workers. First, municipalities can attract skilled and highly-educated foreign workers by preserving their historic districts, and developing them in order to maintain their historic and leisurely value. Moreover, since the share of migrants is also found to be an important factor for location decisions, urban policy can be aimed to stress the positive effects and minimize the possible negative effects which may result from it. Although an existing community of migrants is not an amenity which can be developed artificially as part of an urban policy, municipalities can still aim to preserve the positive effects from an existing community. One possible example of such policy may be encouraging language proficiency courses, which may reduce communication barriers.

Our research's findings contribute to the literature by identifying housing preferences and valuation of urban amenities of skilled and migrant workers based on their individual characteristics. We addressed challenges such as model endogeneity, spatial dependence and omitted variable bias by estimating a two-step residential sorting model and by instrumenting for housing prices and the existing share of migrants. We find that social interaction and labor market amenities are important determinants of location decisions, and that the location preferences of skilled migrants show more similarities to those of skilled native workers. If data availability permits, further research on the valuation of urban amenities by population groups should also focus on the country of origin composition of the group of migrants. This will be useful in identification of individual characteristics, and would allow including the level of cultural diversity within municipalities into the model by using a diversity measure.

7 References

Adamson, D. W., Clark, D. E., & Partridge, M. D. (2004). Do Urban Agglomeration Effects and Household Amenities have a Skill Bias?*. *Journal of Regional Science*, *44*(2), 201–224.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. *Operational Regional Science Series* (p. 289).

Anselin, L., Bera, A., Florax, R., & Yoon, M. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban …*, *26*, 77–104.

Åslund, O. (2005). Now and forever? Initial and subsequent location choices of immigrants. *Regional Science and Urban Economics*, *35*(2), 141–165.

Bartel, a P. (1989). Where do the new U.S. immigrants live? *Journal of Labor Economics*, *7*(4), 371–91.

Bauer, T., Epstein, G. S., Gang, I. N., & Al, E. T. (2007). The Influence Of Stocks and Flows On Migrants' Location Choices. *Research in Labor Economics*, *26*(06), 199–229.

Bayer, P., Ferreira, F., & Mcmillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, *115*(4), 588–638.

Bayer, P., Mcmillan, R., & Rueben, K. (2004). An equilibrium sorting model of sorting in an urban housing market. *NBER*.

Bellini, E., Ottaviano, G., Pinelli, D., & Prarolo, G. (2008). Cultural diversity and economic performance: evidence from European regions. *Hamburg Institute of International Economics (HWWI)*, (3-14).

Berndt, C. A. E. K., Hall, B. H., Hall, R. E., & Hausman, J. A. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, *3*(4), 653–665.

Berry, S., Levinsohn, J., Pakes, A., & Berry, B. Y. S. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, *63*(4), 841–890.

Borjas, G. J. (1994). Ethnicity, neighborhoods, and human capital externalities. *NBER*, (No. w4912).

Brown, W. M., & Scott, D. M. (2012). Human Capital Location Choice: Accounting for Amenities and Thick Labor Markets. *Journal of Regional Science*, *52*(5), 787–808.

Brueckner, J., Thisse, J., & Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review*, *43*, 91–107.

Damm, A. P. (2012). Neighborhood Quality and Labor Market Outcomes: Evidence from Quasi-Random Neighborhood Assignment of Immigrants. *CReAM Discussion Paper*, (35).

Drukker, D. M., Egger, P., & Prucha, I. R. (2013). On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors. *Econometric Reviews*, *32*(5-6), 686–733.

Drukker, D. M., & Prucha, I. R. (2011). A command for estimating spatial-autoregressive models with spatial-autoregressive disturbances and additional endogenous variables. *The Stata Journal*, (1), 1–13.

Eeckhout, J., Pinheiro, R., & Schmidheiny, K. (2010). Spatial sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled.

Florida, R. (2003). Cities and the Creative Class. *City & Community*, *2*(March), 3–19.

Glaeser, E., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, *1*(1), 27–50.

Glaeser, E., & Resseger, M. G. (2010). the Complementarity Between Cities and Skills. *Journal of Regional Science*, *50*(1), 221–244.

Glaeser, E., & Saiz, A. (2003). The rise of the skilled city. *Harvard Institute of Economic Research (HIER)*, (2025).

Goldin, C., & Katz, L. (2009). *The race between education and technology* (Vol. 12984).

Gottlieb, P. D., & Joseph, G. (2006). College-To-Work Migration of Technology Graduates and Holders of Doctorates Within the United States. *Journal of Regional Science*, *46*(4), 627–659.

Hunt, J., & Gauthier-Loiselle, M. (2008). How much does immigration boost innovation?

Jaeger, D. A. (2006). Green Cards and the Location Choices of Immigrants in the United States, 1971-2000. *Institute for the Study of Labor (IZA)*, (IZA Discussion Papers, No. 2145).

Klaiber, H. A., & Phaneuf, D. J. (2010). Valuing Open Space in a Residential Sorting Model of the Twin Cities. *Journal of Environmental Economics and Management*, *60*(2), 57–77.

Marlet, G., & van Woerkens, C. (2005). Tolerance, aesthetics, amenities or jobs? Dutch city attraction to the creative class. *05*. Utrecht University.

McFadden, D., & Train, K. (2000). Mixed MNL Models For Discrete Response. *Journal of Applied Econometrics*, *15*(5), 447–470.

Mocetti, S., & Porello, C. (2010). How does immigration affect native internal mobility? New evidence from Italy. *Regional Science and Urban Economics*, *40*(6), 427–439.

Moretti, E. (2004). Human capital externalities in cities. *Handbook of Regional and Urban Economics*, *4*, 2243–2291.

Moretti, E. (2010). Local Multipliers. *American Economic Review*, *100*(2), 373–377.

Moretti, E. (2012). *The New Geography of Jobs*. *Journal of Economic Geography*. Houghton Mifflin Harcourt.

Moretti, E., & Thulin, P. (2013). Local multipliers and human capital in the United States and Sweden. *Industrial and Corporate Change*, *22*(1), 339–362.

Munshi, K. (2003). Networks in the modern economy: mexican migrants in the u. s. labor market. *Quarterly Journal of Economics*, (May), 549–598.

Niebuhr, A. (2010). Migration and innovation: Does cultural diversity matter for regional R&D activity? *Papers in Regional Science*, *89*(3), 563–585.

Ottaviano, G., & Peri, G. (2006). The economic value of cultural diversity: evidence from US cities. *Journal of Economic Geography*, *6*(1), 9–44.

Ozgen, C., Nijkamp, P., & Poot, J. (2011). Immigration and Innovation in European Regions.

Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, *90*(6), 1257–1278.

Rodríguez-Pose, A., & Ketterer, T. D. (2012). Do Local Amenities Affect the Appeal of Regions in Europe for Migrants?*. *Journal of Regional Science*, *52*(4), 535–561.

Suedekum, J., Wolf, K., & Blien, U. (2009). Cultural diversity and local labour markets. *Regional Studies*.

Van Duijn, M., & Rouwendal, J. (2013). Cultural heritage and the location choice of Dutch households in a residential sorting model. *Journal of Economic Geography*, *13*(3), 473–500.

Appendix A: Randstad Municipalities: Full characteristics table *(Source: CBS, NVM, RCE)*

#	Municipality	COROP area (NUTS3)*	Housing prices	P.migrants	Accessibility (av. Dist to train station)	Jobs	Monuments	Nature	LQ (ICT industries)
1	's-Gravenhage	1	200,282	47.3	3	268.86	1154	11.7	3.47
2	Aalsmeer	4	239,541	14.6	8.1	16.82	36	3.4	1.79
3	Alblasserdam	10	193,780	12.7	9.8	8.42	22	2.1	0.48
4	Albrandswaard	11	201,689	16.1	8.7	7.95	29	10.4	0.45
5	Almere	6	162,833	36.4	1.8	67.23	4	24	0.93
6	Alphen aan den Rijn	7	211,606	19.8	2.6	32.89	74	0.1	0.93
7	Amersfoort	3	219,143	22.4	2.4	85.5	424	7.8	2.1
8	Amstelveen	4	283,645	32.8	6.2	43.39	44	9.3	1.79
9	Amsterdam	4	312,538	49.5	2.5	513.18	7442	2.2	1.79
10	Baarn	3	263,941	15.5	1.5	11.81	151	50.7	2.1
11	Barendrecht	11	201,890	18.6	3.1	21.52	20	2.9	0.45
12	Beemster	4	223,123	8.2	6.5	3.04	90	0.2	1.79
13	Bergambacht	7	216,746	7.1	12.5	3.68	33	1.5	0.93
14	Bernisse	11	193,200	8	16.4	2.6	95	1.4	0.45
15	Beverwijk	13	187,115	21.4	1.7	20.72	35	11.7	0.22
16	Binnenmaas	11	199,482	8.1	9.9	7.25	73	2.5	0.45
17	Blaricum	2	321,091	16.5	7.6	1.87	69	22.4	2.43
18	Bloemendaal	5	366,400	17	1.8	4.16	237	66	1.44
19	Bodegraven	7	219,362	11.9	1.5	8.94	27	0.6	0.93
20	Boskoop	7	210,972	13.3	1.9	4.01	12	0.1	0.93
21	Brielle	11	214,160	11.6	11.6	4.73	374	2	0.45
22	Bunnik	3	255,884	10.9	2.5	7.05	56	4.7	2.1
23	Bunschoten	3	214,160	7.1	8.9	9.81	14	0.4	2.1

24	Bussum	2	268,466	19.4	1.4	11.9	49	12.5	2.43
25	Capelle aan den IJssel	11	199,285	29.5	3.2	45.54	17	0.4	0.45
26	Castricum	13	237,395	9.6	2.5	8.8	48	28.3	0.22
27	Cromstrijen	11	196,120	6.4	18.6	4.05	14	2.1	0.45
28	De Bilt	3	270,622	14.9	2.4	15.88	160	23.4	2.1
29	De Ronde Venen	3	248,570	12.6	10.6	11.86	35	0.3	2.1
30	Delft	8	230,609	29.7	1.5	52.49	689	4.9	0.65
31	Diemen	4	255,118	37	1.1	33.32	16	22.7	1.79
32	Dirksland	11	171,182	3.9	34.1	2.78	31	18.9	0.45
33	Dordrecht	10	179,240	27.2	1.7	57.43	891	14	0.48
34	Edam-Volendam	4	239,541	7	9	11.11	179	0	1.79
35	Eemnes	3	253,338	12.2	6.8	2.09	43	3.3	2.1
36	Giessenlanden	10	218,050	5.7	4.1	3.81	65	0.6	0.48
37	Goedereede	11	213,092	4	29.2	4.21	137	34.5	0.45
38	Gorinchem	10	204,532	23.3	1.8	23.77	218	1.9	0.48
39	Gouda	7	198,091	22.1	1.6	37.74	354	0.9	0.93
40	Graafstroom	10	190,133	3.6	7.5	3.5	91	1.7	0.48
41	Graft-De Rijp	4	198,289	7.2	11.2	1.47	137	1.9	1.79
42	Haarlem	5	251,319	24.5	2	70.58	1182	1.4	1.44
43	Haarlemmerliede en Spaarnwoude	5	238,346	12.3	5.2	1.4	16	13.4	1.44
44	Haarlemmermeer	4	225,819	22.3	4.3	149.43	20	0.5	1.79
45	Hardinxveld-Giessendam	10	211,396	4.6	2	8.42	11	3.5	0.48
46	Heemskerk	13	208,664	17.5	2.9	8.17	22	38.4	0.22
47	Heemstede	5	328,890	17.2	2.3	7.6	97	15.1	1.44
48	Hellevoetsluis	11	183,045	16.7	16.8	9.93	52	6	0.45
49	Hendrik-Ido-Ambacht	10	201,487	11.6	4.1	7.37	9	1.1	0.48
50	Hillegom	9	226,495	13.9	2.7	7.2	7	1.5	0.55
51	Hilversum	2	242,433	22	1.6	53.81	211	42.5	2.43
52	Houten	3	219,582	12.9	2.2	19.84	134	1.3	2.1
53	Huizen	2	260,532	18.6	7.5	12.86	39	29.3	2.43
54	IJsselstein	3	204,532	19.1	11.3	11.26	68	1.5	2.1
55	Kaag En Braassem	9	311,001	7.1	7.7	6.88	60	0	0.55
56	Katwijk	9	251,319	8.5	6.5	24.62	51	17.4	0.55
57	Korendijk	11	196,316	5.5	22.8	1.77	46	6	0.45
58	Krimpen aan den IJssel	11	215,665	12.3	7.4	8.8	14	2.9	0.45
59	Landsmeer	4	267,930	12.5	8	1.91	7	12.8	1.79
60	Lansingerland	11	216,313	13.7	6.7	18.57	17	3	0.45
61	Laren	2	361,667	20.9	5.1	5.48	98	50.3	2.43
62	Leerdam	10	204,737	21	1.6	8.75	54	3.5	0.48
63	Leiden	9	251,571	27.2	1.5	67.88	1243	1	0.55
64	Leiderdorp	9	244,136	19	3.8	10.01	32	1.4	0.55
65	Leidschendam-Voorburg	1	220,903	26.9	2.2	28.53	113	4.1	3.47
66	Leusden	3	230,840	12.8	6.1	14.22	53	34.2	2.1
67	Liesveld	10	211,817	4.9	13.4	4.37	90	1.9	0.48
68	Lisse	9	231,765	11.5	6.8	9.8	41	9.4	0.55
69	Lopik	3	208,457	6.4	15	4.2	127	2.1	2.1
70	Maassluis	11	215,019	23.3	1.3	7.53	41	3.2	0.45
71	Middelharnis	11	190,895	5.4	37.1	7.4	159	1.7	0.45
72	Midden-Delfland	8	254,353	8.2	3.4	5.53	99	1.8	0.65
73	Montfoort	3	227,176	8.2	6.8	4.85	98	1.6	2.1
74	Muiden	2	276,642	15.7	5.2	1.54	75	7.9	2.43
75	Naarden	2	296,998	18.4	2	10.67	163	29.3	2.43
76	Nederlek	11	207,208	7.7	11.6	3.92	24	5.8	0.45
77	Nieuw-Lekkerland	10	200,883	5.9	12.8	2.2	26	10	0.48
78	Nieuwegein	3	206,175	22.4	7.5	51.22	80	3.3	2.1
79	Nieuwkoop	7	227,403	7.8	10.3	7.89	27	7.6	0.93
80	Noordwijk	9	278,307	14.8	5	11.35	74	52.6	0.55
81	Noordwijkerhout	9	238,823	12	5.5	6.48	12	8.3	0.55
82	Oegstgeest	9	275,813	21.4	3.1	6.34	44	3.7	0.55
83	Oostflakkee	11	156,136	5.1	31	2.02	28	7	0.45
84	Oostzaan	4	251,071	10.8	5.4	2.67	7	11	1.79

85	Oud-Beijerland	11	212,879	9.2	14.9	12.13	28	0.5	0.45
86	Ouder-Amstel	4	287,932	20.2	3.2	9.3	36	0.9	1.79
87	Ouderkerk	11	206,381	6	8.9	2.62	38	1.1	0.45
88	Oudewater	3	245,605	6.3	9.4	3.1	159	0.1	2.1
89	Papendrecht	10	197,300	15	6.5	11.94	2	4.1	0.48
90	Pijnacker-Nootdorp	1	224,244	15.5	4	12.21	22	7.3	3.47
91	Purmerend	4	196,709	23.2	1.9	26.44	30	12.1	1.79
92	Renswoude	3	224,915	4.6	4.6	1.82	34	5.5	2.1
93	Rhenen	3	216,963	9.7	2.6	5.03	53	32.9	2.1
94	Ridderkerk	11	200,482	14.4	5.6	20.87	57	2.1	0.45
95	Rijnwoude	7	218,705	7.8	5.8	6.53	81	0.7	0.93
96	Rijswijk	1	197,300	27.8	1.8	35.24	76	0.4	3.47
97	Rotterdam	11	182,313	46.9	2.8	370.03	466	2.9	0.45
98	Schiedam	11	164,141	33.5	1.7	35.26	238	1.7	0.45
99	Schoonhoven	7	217,832	15	12.9	4.82	148	1.2	0.93
100	Sliedrecht	10	205,147	11.3	1.9	13.57	8	0.6	0.48
101	Soest	3	243,161	19.1	2.2	19.83	43	38.3	2.1
102	Spijkenisse	11	172,384	22.7	13.7	21.36	5	5.2	0.45
103	Stichtse Vecht	3	248,570	11.9	2.9	6	196	1.5	2.1
104	Strijen	11	211,183	7.5	13.3	2.23	14	3.5	0.45
105	Teylingen	9	245,851	12.7	2.7	12.53	107	2.3	0.55
106	Uitgeest	13	204,123	10.3	1.5	3.76	20	1.7	0.22
107	Uithoorn	4	228,086	18.4	11.5	12.56	12	0.4	1.79
108	Utrecht	3	250,316	31.5	1.9	229.85	1459	2	2.1
109	Utrechtse Heuvelrug	3	259,232	13.5	4.4	20.98	421	45	2.1
110	Veenendaal	3	217,397	15.8	1.4	30.25	16	3.4	2.1
111	Velsen	13	227,858	15.2	2.5	31.46	138	27.6	0.22
112	Vianen	3	205,557	14.3	10.4	12.07	181	8.1	2.1
113	Vlaardingen	11	184,514	24.4	1.9	23.3	56	11	0.45
114	Vlist	7	231,997	5.7	6.1	3.14	91	0.1	0.93
115	Voorschoten	9	263,413	20.5	1.5	5	103	6.4	0.55
116	Waddinxveen	7	208,875	12.8	1.2	12.14	6	0.9	0.93
117	Wassenaar	1	329,220	29.8	3.9	7.49	272	49.3	3.47
118	Waterland	4	254,353	10.3	10.3	3.15	317	2.3	1.79
119	Weesp	2	260,532	23.7	1.3	10.25	221	0.5	2.43
120	Westland	8	230,840	9.5	6.5	57.14	76	4.8	0.65
121	Westvoorne	11	246,836	9.1	17.6	3.62	29	28.3	0.45
122	Wijdemeren	2	282,513	11.4	6	6.34	249	14.7	2.43
123	Wijk bij Duurstede	3	233,160	10.5	10.5	5.09	166	8.5	2.1
124	Woerden	3	230,379	12.5	2.9	25.58	103	0.7	2.1
125	Wormerland	12	207,831	11	2.5	4.09	33	1.1	0.25
126	Woudenberg	3	240,021	7.7	5	4.05	30	24.5	2.1
127	Zaanstad	12	192,621	25.7	2.1	56.62	258	5.6	0.25
128	Zandvoort	5	277,473	18.7	1	4.66	11	77.9	1.44
129	Zederik	10	220,024	4	8.4	4.47	106	3.6	0.48
130	Zeevang	4	205,969	9.5	8.2	1.25	23	1	1.79
131	Zeist	3	264,999	21.4	3.9	36.21	155	44.9	2.1
132	Zoetermeer	1	200,482	27.7	3.5	48.9	16	1.3	3.47
133	Zoeterwoude	9	245,605	9.3	4	7	40	0.7	0.55
134	Zuidplas	11	210,130	13.2	1.2	6.5	14	4.8	0.45
135	Zwijndrecht	10	192,236	18.4	2.3	19.52	13	1.5	0.48

*COROP area (NUTS3)	Number
Agglomeratie S-Gravenhage	1
Het Gooi En Vechtstreek	2
Utrecht	3
Groot-Amsterdam	4
Agglomeratie Haarlem	5
Flevoland	6
Oost-Zuid-Holland	7
Delft En Westland	8
Agglomeratie Leiden En Bollenstreek	9
Zuidoost-Zuid-Holland	10
Groot-Rijnmond	11
Zaanstreek	12
Ijmond	13

Appendix B: Map of Randstad municipalities alternative specific constants