Understanding the Impact of Geographical Context on Subjective Well-Being of Urban Citizens

Afra Mashhadi
Bell Labs
Nokia
afra.mashhadi@bell-labs.com

Sourav Bhattacharya
Bell Labs
Nokia
sourav.bhattacharya@bell-labs.com

Fahim Kawsar
Bell Labs
Nokia
fahim.kawsar@bell-labs.com

ABSTRACT

Over the past years, the impact of spatial characteristics on subjective well-being has started to receive attention but mostly on the macro granularity of sub-national level. In addition the studies that focused on the spatial scale of urban cities and their neighbourhood have mainly examined the influence of environmental perspectives, land use and urban morphological features. The influence of geographical contexts such as city attributes however have been studied sporadically with somewhat contradicting observations. In this work we focus on the theoretical foundation of subjective well-being and through the discrepancy theory we examine the impact of what is offered and desired by citizens on their subjective well-being. We model functionalities a neighbourhood offers in terms of density, diversity and rarity of its human-made amenities. To infer whether these functionalities fit the desire of residents we model their propensity to travel in and out of an area through large scale analysis of urban mobility flows. Our analysis supports discrepancy theory by showing that the the gap between what is offered and desired is a good predictor for subjective well-being.

CCS Concepts

•**Information systems** → **Geographic information systems;** •**Human-centered computing** → *Social engineering (social sciences); Ubiquitous computing;* Empirical studies in collaborative and social computing;

Keywords

Subjective Well-Being; Urban Mobility; Data Driven Urban Planning

1. INTRODUCTION

Subjective Well-Being (SWB), also commonly referred to as *happiness*, has recently become a focal research point by many different disciplines. Indeed, understanding people's SWB which was traditionally a dilemma examined by philosophers, has also become an examined field in economy and geography in the recent years. In economy, the discipline of "happiness economics", which predominantly accounts for the hedonic attribute, has emerged and strives to find out how well the economic output data such as GDP can represent SWB [11, 14]. As part of this trend, many governing bodies including the European Union has included SWB indicator in their traditional household surveys (i.e., Census) which had previously only captured the classical economics attributes such as unemployment rate and household salary.

Despite the received attention from the policy makers and researchers of various fields, a dimension of SWB that has been less investigated is in regards with the impact of geographical contexts [31, 3, 2]. That is the effect that urban physical structure, social environment and neighbourhood facilities may have on citizens' SWB. Indeed as urban cities rapidly growing to host millions more people across the planet, it becomes increasingly important to understand what *geographical contextual factors* contribute to the subjective well-being of them and their citizens.

In this vein, small body of literature has examined the impact of geographical context on SWB at national level [10, 2, 13]. These works have mainly reported of positive association between SWB and those of economic development (e.g., GDP), cultural differences (e.g., individualism vs collectivism) and climate factors (e.g., pollution and weather). Going from macro-level to micro-level granularity of sub national and urban cities, even smaller body of literature with somehow inconsistent findings exists. Artifwidodo et. al. [1] reported that living in urban centre is associated with higher life satisfaction, while authors in [5, 26] claimed to the contrary. Lowless et al. [18] has shown the negative effect of commute time on SWB. Brereton et al.[6] reported that walking distance to public transport has negative impact on SWB. Dittmann et al. [12] showed that social cohesiveness in terms of visits from/to the area has positive effects on SWB.

With the advent of user generated content on Web 2.0 platforms (e.g. Foursquare etc.), computational social scientist have also examined different facets of urban cities. Quercia et al. leveraged the static images of a city captured from Google street view and annotated by anonymous crowd to build a happiness map of a city [24]. Similarly in [22], the authors examined the features that contribute to the walkability of a city based on user generated content. Zambaldi et al. proposed a new web image ranking technique to identify memorable city pictures based on the prediction of whether a neighbourhood makes people happy [33]. In [25, 23] authors used Twitter data to infer citizens' happiness by examining the conversation topics and sentiments respectively.

We take the above works one step further by investigating the impact that geographical attributes have on the subjective well-being of citizens. In order to do so, we examine SWB from the theoretical perspective of *discrepancy theory* that states the SWB is influenced by whether a place fits citizens needs [20]. To examine this theory, we extract features from urban city which correspond to what amenities citizens have on their neighbourhood and whether these amenities suffice their needs. To this end, we study the subjective well-being of the Greater London's neighbourhood in micro granularity. Firstly by characterising the functional properties that each area offers in terms of human-made amenities. We do so by leveraging a map dataset to extract features that are grounded on Information Retrieval field and best capture the diversity, diversity and rarity of the amenities. Secondly we model citizen's need through their propensity to travel in and out an area by modelling citizens urban mobility. For which we exploit graph theory metrics to extract urban mobility features from a transit dataset originating from London's automated fare collection system.

Our results indicate that a multi-variant regression model based on the discrepancy theory is a good predictor of the SWB. From a practical standpoint, Our results suggest possibilities for design of tools and systems that can help urban designers to model the city in an informed manner, design cityscapes that capture happiness.

2. TERMINOLOGY AND DEFINITION

Previous literature has classified well-being based on multiple dimensions [9, 29, 31]. First and most dominant classification is subjective well-being versus objective welfare, where the former describes the individual's *perception* of how one lives as opposed to the latter which is concerned with a series of objective indicators of welfare. The second dimension is based on the philosophical perspective of hedonic versus eudaimonic. The hedonic approaches believe that well-being is a measure of pleasure attainment and pain avoidance, whereas eudaimonic is focused with the meaning of life and describes happiness as a goal oriented activity and a measure of self-realisation. Finally another classification of well-being is the distinction between external and internal qualities of life. That is whether the well-being is caused by the environmental conditions (external) or by personal abilities or psychological outcomes (internal).

In this paper we focus on the definition of SWB as a measure of *'subjective'*, *'hedonic'* and *'external'* well-being. This definition allows us to examine the impact of geographical context on SWB without being concerned with eudaimonic aspects which are deeply rooted in psychological well-being. Although we focus on the 'subjective' dimension of well-being, we will also account for objective welfare as previous literature has suggested that socio-economic factors such as income have direct impact on the SWB [15].

2.1 Measurement

SWB is frequently measured based on two dimensions of *cognitive* and *affective*. The cognitive measure of SWB corresponds to the general cognitive evaluation of ones satisfaction with her various domains of life. This measure is often referred to as "life satisfaction" and in surveys is usually measured by the question: *"All things considered, how satisfied are you with your life as a whole these days?"*. The affective measure denotes the happiness component of SWB and is generally measured by the question: *"Generally speaking, how happy are you these days?"*. Subjective well-being is often computed by the composite indexes of both cognitive and affective measures.

In this paper we use a measure of SWB index that has been collected by the UK's Office for National Statistics and consists of four indicators of *life satisfaction*, *worthwhileness*, *anxiety*, and *happiness*. These indicators have been captured based on the four following household survey questions:

- Overall, how satisfied are you with your life nowadays?
- Overall, to what extent do you feel the things you do in your life are worthwhile?
- Overall, how happy did you feel yesterday?
- Overall, how anxious did you feel yesterday?

While the first two questions capture the cognitive measures (life satisfaction) the third and fourth aim to capture affective measures (happiness). The collected data is then turned into a composite Subjective Well-being (SWB) score, in which each indicator contributes equally. A positive SWB value indicates a high level of satisfaction with life and positive affect (or in simpler terms, happy). A negative SWB value on the other hand represents negative affect towards life and low life satisfaction. We select this measure of SWB as it is available in micro granularity required for our study as we will describe in the next section.

2.2 Theoretical Framework

Various theoretical perspectives exists that aim to explain the reasons behind people's subjective well-being. In this paper we borrow from the *discrepancy* theory. The discrepancy theory argues that happiness is function of perceived discrepancies between what one has and what she desires [20]. Applied to the spatial context, the discrepancy theory implies that one's subjective well-being is influenced by whether the place fits his/her needs or aspirations [16].

Based on this theory we posit the following research question: *is the gap between what is offered in a neighbourhood and what is desired by the residents a good predictor of the residents subjective well-being?* That is for discrepancy theory to hold we expect to observe an inverse correlation between the gap between 'what is offered' and 'what is desired' and subjective well-being. In other words, when this gap increases the subjective well-being of the residents should decrease.

To answer this question, we first need to capture 'what is offered' by neighbourhood and whether it is 'what is desired' by residents. We capture the former element by the functionalities a neighbourhood affords. We define these affordances as the number of amenities an area offers as well as the diversity of their functionalities (e.g., cafes, hospitals etc). We capture the latter element of 'what is desired' through the propensity to travel into and outside of an area. That is if the urban offering fits the residents desires, we expect to observe citizens from other parts of city to be also drawn to the area (i.e., appeal of an area). Inversely, if the place does not fit residents desire we expect to observe the residents travel from their current neighbourhood for alternative offerings in the city. In the next section we describe the features that correspond to the components of this theory and the datasets from which we extract these features.

3. METHOD AND DATASET

To answer the posited research question, we employ a correlation and regression analysis. In doing so we use SWB as the dependent variable and features corresponding to 'what is offered' and 'what is desired' as independent variables. To extract these variables we require three datasets one corresponding to fine grain SWB information of city, another to capture the functional facets of the city and finally a third dataset that presents mobility flow of the citizens so to capture the propensity to travel in and out of areas. For this purpose we have chosen to study our hypothesis on the city of London where all the required datasets are available from the same time period (2010) and in a fine geographical granularity of wards (i.e., official UK administrative boundaries). In the remaining of this section we describe the datasets at hand that we used in order to validate the proposed hypothesis and the extracted features (independent variables).

3.1 Subjective Well-Being

Presently, London is composed of 649 wards, and in 2010, SWB scores were available for 624 wards out of these 649 wards [19]. Figure 1a illustrates the distribution of SWB for the city of London at ward scale, where the red areas are indicative of lower SWB and green areas represent higher SWB values (i.e., higher happiness). The grey shaded wards are those for which the SWB score was unavailable.

In addition to SWB, the UK National Office for Statistics makes available data regarding the Indices of Multiple Deprivation (IMD). IMD is a measure of socio-economic deprivation that is published in small geographic granularity known as Lower-layer Super Output Areas (LSOA) in England. The IMD consists of seven domain indicators corresponding to the *depravation of*: income, employment, health & disability, education & skills, barriers to housing & services, living environment, and crime. Research in social economics have previously shown that the SWB is often impacted by factors related to the general well being such as income, and health [11, 28]. To this end, past studies in economics [11, 8, 28] highlighted that the relationship between income and happiness as linear-log relation. These studies also argued that once "basic needs" have been met, higher income may no longer be highly associated with higher subjective well-being. That is while income is closely tied to happiness, there are also smaller measures of happiness. Figure 1c presents the distribution of the mean IMD score at ward levels in Greater London, where the lower scores (i.e., green areas) correspond to the wealthier areas in terms of socio-economic standings. We can observe that some areas are high in terms of wealthiness but have low SWB scores. To understand this relationship better we have computed the correlation between the SWB and the composite IMD along with each of its seven indicators. Figure 2 presents the significant result of the Pearson correlation analysis ($p - value < 0.09$). We can observe that the composite IMD value shows a weak negative correlation with the SWB, i.e., the *lack* of depravation is weakly correlated with the subjective well-being. When we take a deeper look into the indicators of IMD, we can see the highest correlation ($r = -0.17$) is observed based on the income deprivation and unemployment ($r = -0.15$) as suggested by [4, 7]. Delving upon these observations, we argue that that SWB is not limited to IMD but also impacted by the geographical contexts (neighbourhood attributes) and the choices available to the citizens.

3.2 Map Data

In order to model the functional facets of the city in terms of 'what is offered' in its constituent neighbourhoods, we require a dataset of London's PoIs corresponding to the same time period as the other datasets at hand (i.e., 2010). For this purpose, we selected Navteq,[1] the leading global provider of maps and location data, covering not only roads but also millions of PoIs of varying nature, from restaurants to hospitals to cash machines. Being a commercial service, Navteq's primary objective is to ensure the highest level of accuracy to its data, i.e., the information contained there is factually correct and up-to-date.

This dataset contains of over 35 thousands PoIs distributed across 56 categories as defined by Navteq. Having categorical data enables us to model the diversity of the PoIs in terms of their offerings such as educational offerings (e.g., schools and universities), gastronomy offerings (e.g, restaurants, cafes) and others. To reduce the dimensionality of this data we abstracted the original 56 categories into 8 super-categories.

Table 1: Properties of the functional features.

	Parameter	μ	Distribution	σ
	Density	1.876e-05		2.37e-05
Functional	Rarity	1.08		0.04
features	Diversity	1.72		0.19

We extract three main features: (i) *PoI density*, (ii) *PoI diversity*, and (iii) *PoI rarity* to capture concentration, variability and rarity of amenities in a neighbourhood. These features enable us to capture 'what is offered' in the neighbourhood as part of the discrepancy theory. Table 1 summarises the statistics of all the extracted functional features.

PoI Density: In order to capture the availability of amenities that are offered in each ward we compute PoI *density* of ward w^i by taking the ratio of the number of PoIs n^i found in w^i to the area of w^i.

PoI Diversity: To model the diversity of amenities in each ward based on the defined categories, we follow an information theoretic approach and compute the Shannon entropy of the distribution of PoIs found within a ward.

PoI Rarity: The final feature extracted from this dataset is the *rarity* of amenities which allows us to model the attractiveness of a neighbourhood based on the unique PoIs it hosts (e.g., Zoo). To do this we consider TF-IDF (term frequency - inverse document frequency) normalisation of PoI distribution within wards.

3.3 Urban Mobility Data

While the previous dataset allows us to form an image of what is offered in the urban areas, it does not describe whether the offering fits the residents needs. One way to address needs/desire is by reasoning on the propensity to travel. That is if an area is desirable in terms of its functionalities we expect to observe propensity to travel into that area from all over the city. Likewise, if an area's offering does not fit its residents needs we expect to observe propensity to travel *out* of area to seek alternatives. For this purpose we used Oyster (RFID) card transit data. We are presently limited to data pertaining to the rail and underground subnetworks, which include a total of 583 stations. The dataset at hand

[1] https://company.here.com/here/

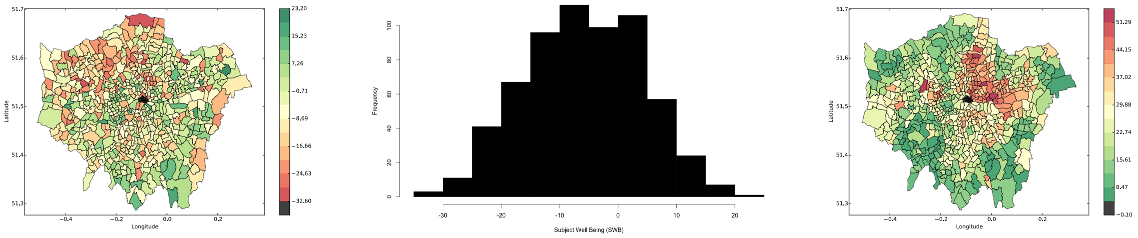

Figure 1: (a) Composite Subjective Well-Being score across various wards, (b) Composite Subjective Well-Being Frequency Distribution, (c) Composite IMD score across various wards in Greater London.

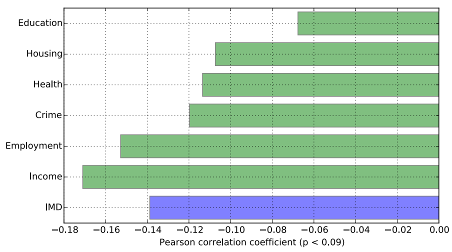

Figure 2: Pearson's correlation between the SWB and the composite and break-down IMD.

consists of a record of every journey taken on the Transport for London network using an Oyster card in the 31 days of March 2010, corresponding to over 66 million journey from 4 million distinct users. The records are anonymised and comprise the touch-in/out station name, and the time of the entrance/exit for each user.

We computed two pre-processing steps on this data. Firstly we accounted for continuous journeys by assuming two trips are part of a same journey if the stopover duration is less than 15 mins. Secondly, we segregated tourists from citizens of London. This segregation is needed when computing correlation analysis with census variables, as they are primarily conducted based on the local house-hold surveys. In so doing, we identify a user as a citizen, if she has used the transportation network on three different weeks (or more), out of the five weeks of Oyster card transit data. The pre-processing steps filtered over 49% of the users and identified over 2 million unique citizens contributing to over 57 million trips.

Forming Transition Matrices

We represent passenger flows among various stations with a directed graph $G = (V, E)$, where V is the set of all stations and the weight of the edge (i, j) between station i and j indicates the number of passengers who made a trip from station i to station j. Internally the graph is represented by a adjacency matrix $\mathcal{M}_{|s|\times|s|}$ (a square matrix), where $|s|$ is the total number of stations in the transportation network. The matrix $\mathcal{M}_{|s|\times|s|}$ is asymmetric in nature, i.e., $\mathcal{M}(i, j) \neq \mathcal{M}(j, i)$. In other words, the number of passengers travelling from station i to j is different from the number of passengers travelling from station j to i.

We then construct a $3D$ transition matrix, where the third dimension represents time as illustrated in Figure 3(a). We divide the entire duration of the Oyster dataset into discrete and non-overlapping time buckets (indexed by t_i in the figure), and compute an adjacency matrix for each time buckets. In this work we consider a constant time bucket-width of 6 hours, i.e., we divide each days into four different time zones: (i) *early* (0-5 hrs), (ii) *morning* (6-11 hrs), (iii) *afternoon* (12-17 hrs), and (iv) *evening* (18-23 hrs). These time zones intuitively correspond to different time dependent flow patterns of passengers in the city (e.g., rush hours).

The $3D$ transition matrix captures accurately passenger flow across stations, however, to correlate the transit features with the SWB of the citizens, we convert the transition matrix for stations into another $3D$ matrix of *wards*. This is done by aggregating all inflow/outflow of passengers from all stations within a ward to any other wards. Note that, after this mapping, intra-ward passenger flows are represented by a self loops, consequently the diagonal element of the ward transition matrix can be non-zero. As the Oyster dataset provides tube and rail network transit information, wards with no stations do not appear in the inter-ward transition matrix, reducing our sample size to 357 wards only.

Lastly, to make distinction between in-flow and out-flow in passenger transitions, we divide each ward transition matrices into their corresponding *upper-* and *lower-triangular* matrices. Each triangular matrix is then converted into a symmetric matrix by taking mirror image of the elements with respect to the leading diagonal. Figure 3(b)-(c) provides an overview of the in-flow/out-flow symmetric matrix generation process.

Feature Extraction

From this processed dataset, we extract two features that correspond to the propensity to travel out and in an area. These features enable us to understand the residents need to seek alternatives and the desirability of the area respectively. To do so, we borrow from graph theory and measure (i) *degree-centrality* and (ii) *eigen-centrality*.

Degree centrality. This feature corresponds to the propensity to travel outside of the immediate ward, which we proxy as the need to seek alternative functionalities from other areas of the city. Given an adjacency matrix $\mathcal{M}$ representing a graph, then the degree $d(i)$ of vertex i is defined as:

$$d(i) = \sum_{j=1}^{n} \mathcal{M}(i, j) \qquad (1)$$

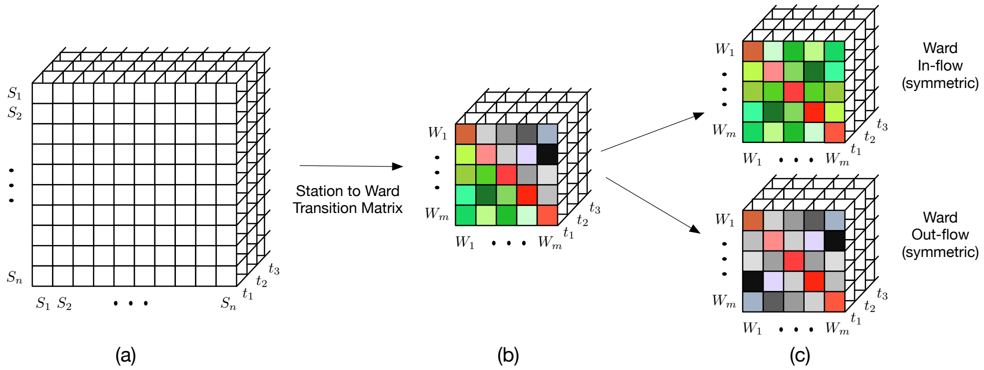

Figure 3: Conceptual 3D visualisation transition matrices of (a)station flow (b)ward flow and, (c)symmetric in-flow/out-flow from a asymmetric ward transition matrix.

We compute degree centralities for out-flow journey for all four different time buckets. We do so by first taking average of all the transition matrices found within each time bucket and then computing the degree centralities. Additionally, we only focus on weekdays, which tend to capture citizens' routine travelling behaviour.

Eigenvector centrality. In order to model appeal of an area, we extract eigenvector centrality [21] of the wards for in-flow inter ward transition matrices for all four time buckets. This metric also allows us to compute influence or importance of a node within the network. Contrary to degree centrality, which gives a simple count statistics of the number of connection a vertex has, eigenvector centrality considers the fact that not all connections within a network are equal. Intuitively, from a social network point of view, eigenvector centrality captures the fact that connections to people who are themselves influential increases the importance of a person. Similarly in the case of urban mobility it captures whether an urban area is a hub for visitors to/from other urban hubs. This feature in essence provides us with an indication of popularity of an urban area. To take into account strengths of connections, the eigenvector centrality $e(i)$ of a vertex i is computed as [21].

$$e(i) = \frac{1}{\lambda} \sum_{j=1}^{n} \mathcal{M}(i,j) \cdot e(j), \tag{2}$$

4. RESULTS

In this section we report the result of our analysis which aims to answer the following research question " is the gap between what is offered in a neighbourhood and what is desired by the residents a good predictor of the residents subjective well-being?" To begin with, we first examine the relationship between the functional attributes of neighbourhood ('what is offered') and their influence on the SWB. We measure the Pearson correlation between the extracted functional features and the SWB across the 624 wards in the Greater London. As some of our parameters are moderately/strongly skewed, we first computed their log-transformation.

Surprisingly, we do not observe any statistically significant correlations between the SWB and the diversity of PoIs. That is having variety of amenities form different categories in a neighbourhood did not exhibit a relation with subjective well-being. Similarly, we did not observe any statistically significant correlation with the PoI density. However we have observed a very weak correlation between SWB and the rarity of the PoIs ($r = 0.1, p - value = 0.05$), that is areas with rare PoIs were more likely to exhibit higher SWB.

However to reaffirm these observations we need to account for the IMD factor as pervious literatures have demonstrated a linear-log relation between SWB and IMD. To account for this factor, we first segmented our data into five subsections based on the IMD, where the top fifth corresponds to the most deprived areas, and the bottom fifth corresponds to the least deprived areas of London. We then repeated our correlation analysis for those *least* deprived wards(df= 69). The result of this analysis reports of an inverse correlation between SWB and PoI density ($r = -0.27$, p-value=0.001). That is the richer wards exhibited a positive relation between the *lack* of amenities and the residents' subjective well-being. Looking deeper into this finding, we observed that in particular lack of PoIs which corresponded to the categories of nightlife ($r = -0.19, p{-}value = 0.04$), and shops ($r = -0.22$, p-value=0.02) contributed to this correlation.

Next, we evaluate the relationship between propensity to travel (in/out) and the functionalities a neighbourhood offers. In so doing, we account for distance from the city centre (as an independent variable) as it has been shown to have a direct effect on trips made by citizens. As before we use the linear regression analysis for which we transform the independent variables (transit features) to logarithmic scale. In the log transformation where the parameters were all negative such as those of Eigen Centrality Inflows, we first transferred them to positive range and added a constant of e^{-8} to ensure they are all non-zero values. As expected we observe a correlation between propensity to travel IN to an area in the *evening* and the rarity of PoIs in that area(Multiple R^2=0.12, p-value<0.001). In addition to rarity, we also observe weak correlation between PoI density and propensity to travel in to an area (Multiple R^2=0.15, p-value<0.001), with a much smaller correlation with propensity to travel out (Multiple R^2=0.05, p-value<0.001) . However we did not observe any relation between propensity to travel and diversity of the PoIs in an area. This perhaps confirms previous studies [34] which have shown that the purpose of travels are often tied to a unique dominant functionality that the destination offers (e.g., shopping district, stadium etc.).

Table 2: Multiple R^2 and p-value of our three models. Significancy. codes: 0 * 0.001 ** 0.01 ***

Model	*Multiple* R^2
Functional Model: SWB $\sim\beta_1$. Density + β_2 . Rarity + IMD	0.03 ***
Propensity Model: SWB $\sim\beta_1$. Eigen Centrality-Inflow + β_2 . Degree Centrality-Outflow + Distance	0.04 *
Final Model: SWB $\sim$All the above features	0.25 **

Brining these observations together, we build three regression models, first based on the functional properties only, second one based on the propensity to travel and a final model based on all the features which corresponds to the discrepancy theory. The first two models exhibit very low multiple R^2 , indicating that the information regarding the functional attributes or propensity to travel by themselves are not a good predictor of SWB. However the final model which leverages both the functional and propensity to travel features is able to predict SWB with a strong level of accuracy (Multiple R^2=0.26). This confirms our initial hypothesis based on discrepancy theory that the gap between what is offered and needed is a good indicator of estimating subjective well-being of the city neighbourhood.

5. RELATED WORK

In an attempt to understand urban cities at a fine level of spatio-temporal granularity, in recent years researchers have focused on computational methods that could automatically profile urban areas and provide insights in terms of their functionalities, depravation and neighbourhood satisfaction (e.g., recognisability, walkability, and happiness).

In [22], the authors examined the urban features that contribute to the walkability of a city based on the social media data of Flickr and Foursquare. Venerandi et al. [30] also leveraged user-generated content (Open Street Map and Foursquare) to profile urban neighbourhoods in terms of functional advantages, which was then used to automatically uncover socio-economic deprivation of urban areas. Zambaldi et al.[33] explored the recognisability of a city by proposing a new image ranking technique that identifies memorable city pictures based on the prediction of whether a neighbourhood makes people happy. In [25, 23] authors used classical text mining techniques to infer citizen happiness from Twitter conversations, and embedded sentiments.

The rich data provided by the smart transport cards has enabled vast literature of research in understanding people's mobility as a mass, as well as uncovering the purpose of the travel by looking at the affordances of the city. In [34], Zhu et. al showed that the affordances and functionalities of the city are major indicators to predict the purpose of a trip. Similarly, however on a larger scale, Yuan and his colleagues studied the functionalities of the city in relation to the mobility trajectories from taxi cab traces, and the PoI data [32]. They proposed a new topic model-based identification technique which takes into account the multi-functionality of an area. More recently, transportation network data has been used to uncover and explain the characteristics of the city such as the socio-economic status. Lathia et al. showed that transport network data (i.e., trip information extracted from automated fare collection systems) can be successfully exploited to infer socio-economic indices of urban neighbourhoods [17]. Smith et al. [27], also used transportation data to identify deprived areas within the Greater London. The authors formulated the identification task as a prediction problem and trained a linear regressor, while considering features from residues of a gravity model, population biases around stations and diversity of user mobilities. They then studied the relation between the derived features and socio-economic status of the areas as captured by Indices of Multiple Deprivation (IMD).

6. CONCLUSION

In this paper we have investigated the impact of geographical context on hedonic, external subjective well-being and have tested it through the discrepancy theory. We have examined features related to what an area offers and what the citizens need, and have shown their possible influence on subjective well-being. We have shown that by bringing these attributes together we can build a regression model that can estimate the subjective well-being of neighbourhoods. From a practical standpoint, the suggested method to estimate the subjective well-being of a neighbourhood can be transformed into a tool that can help urban designers to model the city in an informed manner, design cityscapes that capture happiness [24] and develop novel services that open up new opportunities for the citizens. From a theoretical standpoint, we have shown how datasets acquired for different purposes could be used to understand the relationship between the mobility and utility aspects of a neighbourhood and the subjective well-being of its residents. Methods like ours offer the potential to scale the research on happiness at a planet scale across cultures.

7. ACKNOWLEDGEMENT

We would like to thank Dr. Mirco Musolesi for his feedback on different stages of this work. Part of this researcher and the datasets used in this article was funded and obtained under EIT ICT European Project TravelDashboard.

References

[1] Arifwidodo, S. D. and Perera, R. (2011). Quality of life and compact development policies in bandung, indonesia. *Applied Research in Quality of Life*, 6(2):159–179.

[2] Aslam, A. and Corrado, L. (2011). The geography of well-being. *Journal of Economic Geography*, page lbr041.

[3] Ballas, D. and Dorling, D. (2013). The geography of happiness. *The Oxford handbook of happiness*, pages 465–481.

[4] Ballas, D. and Tranmer, M. (2011). Happy people or happy places? a multilevel modeling approach to the analysis of happiness and well-being. *International Regional Science Review*, page 0160017611403737.

[5] Berry, B. J. and Okulicz-Kozaryn, A. (2011). An urban-rural happiness gradient. *Urban geography*, 32(6):871–883.

[6] Brereton, F., Bullock, C., Clinch, J. P., and Scott, M. (2011). Rural change and individual well-being the case

of ireland and rural quality of life. *European Urban and Regional Studies*, 18(2):203–227.

[7] Clark, A. E. (2003). Unemployment as a social norm: Psychological evidence from panel data. *Journal of labor economics*, 21(2):323–351.

[8] Clark, A. E., Frijters, P., and Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic literature*, pages 95–144.

[9] Diener, E. (2009). *Culture and well-being: The collected works of Ed Diener*, volume 38. Springer Science & Business Media.

[10] Diener, E., Diener, M., and Diener, C. (1995). Factors predicting the subjective well-being of nations. *Journal of personality and social psychology*, 69(5):851.

[11] Diener, E. and Seligman, M. E. (2004). Beyond money toward an economy of well-being. *Psychological science in the public interest*, 5(1):1–31.

[12] Dittmann, J. and Goebel, J. (2010). Your house, your car, your education: The socioeconomic situation of the neighborhood and its impact on life satisfaction in germany. *Social Indicators Research*, 96(3):497–513.

[13] Fischer, R. and Van de Vliert, E. (2011). Does climate undermine subjective well-being? a 58-nation study. *Personality and Social Psychology Bulletin*, 37(8):1031–1041.

[14] Frey, B. S. and Stutzer, A. (2010). *Happiness and economics: How the economy and institutions affect human well-being*. Princeton University Press.

[15] Helliwell, J. F. (2003). How's life? combining individual and national variables to explain subjective well-being. *Economic Modelling*, 20(2):331–360.

[16] Kahana, E., Lovegreen, L., Kahana, B., and Kahana, M. (2003). Person, environment, and person-environment fit as influences on residential satisfaction of elders. *Environment and Behavior*, 35(3):434–453.

[17] Lathia, N., Quercia, D., and Crowcroft, J. (2012). The hidden image of the city: sensing community well-being from urban mobility. In *Pervasive computing*, pages 91–98. Springer.

[18] Lawless, N. M. and Lucas, R. E. (2011). Predictors of regional well-being: A county level analysis. *Social Indicators Research*, 101(3):341–357.

[19] London.gov. London data repository. http://data.london.gov.uk/dataset/london-ward-well-being-scores.

[20] Michalos, A. C. (1985). Multiple discrepancies theory (mdt). *Social indicators research*, 16(4):347–413.

[21] Newman, M. E. (2008). The mathematics of networks. *The new palgrave encyclopedia of economics.*

[22] Quercia, D., Aiello, L. M., Schifanella, R., and Davies, A. (2015). The digital life of walkable streets. In *Proceedings of the 24th International Conference on World Wide Web*, pages 875–884. International World Wide Web Conferences Steering Committee.

[23] Quercia, D., Ellis, J., Capra, L., and Crowcroft, J. (2012a). Tracking gross community happiness from tweets. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 965–968. ACM.

[24] Quercia, D., O'Hare, N. K., and Cramer, H. (2014). Aesthetic capital: What makes london look beautiful, quiet, and happy? In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '14, pages 945–955, New York, NY, USA. ACM.

[25] Quercia, D., Séaghdha, D. Ó., and Crowcroft, J. (2012b). Talk of the city: Our tweets, our community happiness. In *ICWSM*.

[26] Rehdanz, K. and Maddison, D. (2008). Local environmental quality and life-satisfaction in germany. *Ecological economics*, 64(4):787–797.

[27] Smith, C., Quercia, D., and Capra, L. (2013). Finger on the pulse: Identifying deprivation using transit flow analysis. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW)*.

[28] Stevenson, B. and Wolfers, J. (2013). Subjective well-being and income: Is there any evidence of satiation? Technical report, National Bureau of Economic Research.

[29] Veenhoven, R. (1996). Developments in satisfaction-research. *Social Indicators Research*, 37(1):1–46.

[30] Venerandi, A., Quattrone, G., Capra, L., Quercia, D., and Saez-Trumper, D. (2015). Measuring urban deprivation from user generated content. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 254–264. ACM.

[31] Wang, F. and Wang, D. (2016). Place, geographical context and subjective well-being: State of art and future directions. In *Mobility, Sociability and Well-being of Urban Living*, pages 189–230. Springer.

[32] Yuan, J., Zheng, Y., and Xie, X. (2012). Discovering regions of different functions in a city using human mobility and pois. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 186–194. ACM.

[33] Zambaldi, V. F., Pesce, J. P., Quercia, D., and Almeida, V. (2014). Lightweight contextual ranking of city pictures: Urban sociology to the rescue. In *ICWSM*.

[34] Zhu, Z., Blanke, U., and Tröster, G. (2014). Inferring travel purpose from crowd-augmented human mobility data. In *Proceedings of the First International Conference on IoT in Urban Space*, pages 44–49. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).