Mental Illness in Ireland: Simulating its geographical prevalence and the role of access to services

Dr. Karyn Morrissey,

Department of Geography and Planning, University of Liverpool, Liverpool, England

Karyn.morrissey@liv.ac.uk

Professor Graham Clarke,

School of Geography, University of Leeds, Leeds, England

g.p.clarke@leeds.ac.uk

Dr. Paul Williamson,

Department of Geography and Planning, University of Liverpool, Liverpool, England

William@liverpool.ac.uk

Ms. Antoinette Daly,

Mental Health Information Systems Unit, Health Research Board, Dublin, Ireland

a.daly@hrb.ie

Dr. Cathal O'Donoghue,

Rural Economic Development Programme, Teagasc, Galway, Ireland

Cathal.odonoghue@teagasc.ie

Mental Health in Ireland: Simulating the geographical prevalence and the role of access to services for depression

Abstract

Ireland has traditionally reported high rates of admissions to acute psychiatric facilities for mental illness in general. However, data limitations mean that there has been no research on the role of access and the proximity on rates of admissions to acute psychiatric facilities. The Simulation Model of the Irish Local Economy (SMILE) produces synthetic small-area level microdata on self-reported rates of depression. The National Psychiatric Inpatient Reporting System (NPIRS) contains spatially referenced admissions data to acute psychiatric services (both public and private) by diagnosis. Combining the NPIRS and SMILE datasets using propensity score matching techniques produces a small area profile of individuals with depression that includes those who have accessed an acute psychiatric facility as well as those who have not. Linking the NPIRS and SMILE datasets allows one to examine the differential characteristics that lead individuals with depression to seek acute psychiatric services and importantly to see if access to these services is a confounding factor. This paper finds that access as measured in terms of road distance has a significant positive impact on individuals with depression using an acute psychiatric facility.

1. Introduction

The spatial organisation of services and the proximity to services for different populations has been shown to influence health services use (Curtis, 2007; Jones et al. 2010; Campbell et al., 2000; Rushton and West, 1999). With regard to psychiatric services, early research by Jarvis (1850) suggested that populations living in closest proximity to psychiatric facilities were more likely to use them than those living further away (cited in Curtis, 2007). Recent research in the USA found that distance to the nearest mental health treatment facility was a significant barrier to receiving any or adequate psychotherapy, whereas distance increased the likelihood of receipt of pharmacotherapy (Pfeiffer et al., 2011), particularly in the case of rural patients (Fortney et al., 1999). Similarly, McCarty et al., (2007) found that variations in geographic accessibility and continuing care service availability produce gaps in treatment and decreased service utilization for patients. Examining admissions to acute psychiatric hospitals for males in New York City, Almog et al., (2006) found that areas closer to acute psychiatric hospitals had higher admission rates. In contrast, when examining New York City and London, Curtis et al., (2006) found an insignificant relationship between access to acute psychiatric hospitals as measured by bed ratios and admissions. Due to mixed nature of the evidence, Curtis (2007) affirms the need for studies that examine the social and economic determinants of mental health service utilisation to take into consideration geographical proximity to care and the nature of local service provision. The purpose of this paper is to explore whether proximity to service provision is an important factor in explaining rates of admissions to acute psychiatric services in Ireland.

2. Psychiatric Care in Ireland

In line with the guidelines of the 1966 Government Commission of Enquiry on Mental Illness in Ireland, the provision of psychiatric care in Ireland has moved from large-scale in-patient based psychiatric care to a community based service. Day hospitals and day centres have been established and community residential places have increased in Ireland (Tedstone-Doherty and Moran, 2009) with the number of inpatients resident in Irish psychiatric facilities falling from 19,801 in 1963 to 2,812 in 2010 (Daly and Walsh, 2011). Recent research by Tedstone-Doherty and Moran (2009) in Ireland provides insight into the spatial (e.g. spatial location and travel distance) and non-spatial factors (e.g. socioeconomic status, health insurance status, and cultural background) determining an individual's use of psychiatric services. In Ireland an individual is only entitled to free public health care if their household income (from whatever source) is below a certain threshold, or if they have a certain illness (not depression) and/or if they are over seventy years of age. Tedstone-Doherty and Moran (2009) found that there once aspatial factors such as income were controlled for, there was no difference in attendance at a GP for mental health problems. On the other hand, Tedstone-Doherty et al., (2007) found considerable geographical variation in the rate of service uptake for high support community-based residential care. Tedstone-Doherty et al., (2007) also noted that the variation in these rates between Health Service Executive (HSE) regions and counties was not associated with the admission rates to acute psychiatric hospitals and units. It is not clear whether differences in rates of community-based residential care are due to geographical variation in admissions policy between HSE regions or whether other issues such as access to these services is a factor.

Previous research on the small area profile of depression in Ireland found that individuals with the highest rates of self-reported depression have poor access to acute psychiatric facilities (Morrissey et al., 2010). Given the role that acute psychiatric facilities still play in treating individuals with depression (Kelleher and O'Brien, 2001) and the limited potential access to these services for individuals with self-reported depression in Ireland, it is important to identify what factors (both aspatial and spatial) determine if an individual with depression seeks acute psychiatric care. Thus research to date arguably indicates that whilst aspatial factors such as income do not influence psychiatric service use, spatial factors (be it institutional or access based) are important in determining the use of psychiatric services in Ireland. Thus, both spatial (proximity) and aspatial (income and age) factors may arguably impact on an individual's choice to access health services for depression in Ireland.

Research to date on the confounding factors associated with admissions to acute psychiatric hospitals has relied on hospital inpatient records (Congdon, 2002). Given the current emphasis in psychiatric care on community based models, the sole use of hospital admissions data to examine patterns of admissions and their confounding factors creates a predictive tool that examines revealed health care demand rather than actual health care need (Harrison et al., 1995; Congdon, 2002; Van de Velde, 2010). Thus, the ideal framework for examining psychiatric care need is one with the ability to predict levels of illness within the broader population (Congdon, 2002). Merging microdata from the National Psychiatric Inpatient Reporting System (NPIRS) on admissions to acute psychiatric facilities for depression with data from SMILE (Simulation Model of the Irish Local Economy) on the prevalence of depression creates a dataset with individual level information on self-reported depression and admissions to acute psychiatric facilities for depression. Such a dataset allows us to establish the relationship between access to acute psychiatric hospital and depression in Ireland, controlling for a host of demographic and socioeconomic characteristics. Propensity score matching (PSM) offers a method of robustly merging these two datasets, thus allowing us to examine the complex relationship between admissions to an acute psychiatric facilities and individual level locational, demographic and socioeconomic factors (Curtis, 2007).

3. Data and Methodology

3.1 National Psychiatric In-patient Reporting System

The Mental Health Information Systems Unit of the Health Research Board (HRB) in Ireland maintains the National Psychiatric In-patient Reporting System (NPIRS). Data on psychiatric admissions are collected from psychiatric in-patient facilities; including psychiatric hospitals, general hospital psychiatric units, private hospitals, children's centres and the Central Mental Hospital in Dublin. The NPIRS includes data on all admissions, discharges, diagnoses (ICD-10), orders of admission (first admission or not), and lengths of stay in psychiatric in-patient facilities in Ireland. The NPIRS also contains demographic, socioeconomic and geographical details for each admission and discharge (Morrissey et al., 2012). Each record was further supplemented by exogenous data from the Health Service Executive (HSE) on the number of day facilities, outpatient facilities and outpatient sessions offered, in 2002, by the Health Boards within which the admission took place. Data from 2002 are used to develop the initial propensity framework as the data from the spatial microsimulation model are also from 2002. However, an updated spatial microsimulation model is currently being developed using the 2011 Census for Ireland and once validated the framework presented in this paper will be used in conjunction with more recent data.

NPIRS is event-based (Morrissey et al., 2013). One person may have several admissions during the course of any given year and each admission is recorded separately. Thus, to use the NPIRS to examine admission rates for certain mental illnesses, a person-based ID is required. The NPIRS lacks such a unique patient ID. A proxy ID was therefore assigned to each case using year of birth, gender, townland (the smallest administrative division in Ireland), ICD-10 diagnosis code, marital status and socioeconomic group. By way of validation it was ensured that if an individual had more than one admission, their admission and discharge dates made chronological sense. Using the methodology described, it was found that at the individual level there were 15,609 admissions in 2002 (Morrissey et al., 2012). Whilst not everyone that has depression requires acute psychiatric care, data from the National Psychiatric In-patient Reporting System (NPIRS) indicates that depressive disorders accounted for the highest proportion of admissions, 29.5%, to acute psychiatric facilities in 2011 (Daly et al., 2012). This corresponds to 2% of the estimated 300,000 individuals in Ireland reporting depression requiring acute inpatient care.

3.2 Simulated Model of the Irish Local Economy - SMILE

In Ireland, the Living in Ireland Survey (LII) contains the requisite data on self-reported rates of depression, along with demographic, socioeconomic, income and other health data for each individual. However, the level of spatial disaggregation in the LII is limited to a regional and an urban/rural classification. Thus, the spatial resolution of this is too coarse to permit detailed data analysis on depression at the required local level (Morrissey et al., 2010). Furthermore, it is not possible to access the raw LII data and its primary sampling unit within a safe-setting in Ireland as it is in other countries such as the UK. Thus, the availability of spatial data in the LII dataset is constrained to the regional and urban/rural classification scale. Consequently, it is not possible to use multi-level models with the LII to examine the impact of geographical location on health (or any other policy issue) without first generating a geographically referenced population dataset. In contrast, the Small Area Population Statistics (SAPS) 2002 contain a variety of demographic and socioeconomic data at the small area level (ED, electoral District), but lack both physical and mental health data. A variety of model-based techniques (often called small area synthetic estimation) exist for combining the more detailed but aspatial information on depression in the LLI with the spatial information from the SAPS to represent and simulate health outcomes at the individual and small area level (Scarbough et al, 2009). These techniques are based on the prediction of single target parameters at the small area level. For example, in the case of income, known predictors of income (i.e., labour force participation, education level, gender and age) are used to predict income levels for each individual in a small area. These methods include synthetic, sample size, dependent, best linear unbiased predictors, and a variety of Bayesian estimators (Ghosh and Rao 1994; Scarbough et al, 2009; Chaudhuri and Ghosh 2011). These methods borrow 'strength' from a standard population (such as a national dataset) to increase the effective sample size for each small area of interest.

In contrast, spatial microsimulation is a data generation process that reweights survey data to match appropriate local spatial characteristics (e.g., age, sex, education level, socioeconomic group) (Morrissey et al., 2013). The resulting outcome is a separate set of survey weights for each estimation area. If, as in the case of SMILE, integer weights are used, the result is a synthetic population with known individual and household level characteristics and spatial location. As the survey data have been reweighted to match multiple local area constraints, spatial microsimulation, in contrast to model based methods of small area estimation, may be used to estimate several different outcomes simultaneously (Smith, Pearce, and Harland 2011). Weighted survey data also have the advantage that they provide, for non-binary variables, an estimate of the full distribution rather than of a single point. For example, the full distribution of income is simulated rather than the mean or median income. Furthermore, given the computational costs of small area estimation both in terms of resources and time, a strong efficiency rationale exists for developing a model that may be used to examine a number of different policy questions. For these reasons, over the last decade, spatial microsimulation techniques have been increasingly used to examine health and health inequalities (Morrissey et al. 2008, 2010; Edwards et al. 2009; Smith, Clarke, and Harland 2009).

The Simulation Model of the Irish Local Economy (SMILE) is a static spatial microsimulation model that uses a combinatorial optimisation technique, simulated annealing, written in the java language, to match (reweight) the 2000 LII survey to the 2002 SAPS (Morrissey et al., 2008). As outlined fully in Morrissey et al., (2008), through the simulated annealing process a micro-level synthetic dataset for the entire population of Ireland is created. The dataset created by SMILE contains demographic, socioeconomic, labour force and income variables at the micro-level for both individuals and family units. However, spatial microsimulation, like any other data fusion or statistical matching technique rests upon the core assumption of conditional independence (D'Orazio, Di Zio, and Scanu 2006). In other words, it assumes that the spatial variability of *all* survey characteristics will be captured through the reweighting of *selected* survey characteristics to local SAPS totals (Morrissey et al., 2013).

To correct for breaches in this assumption the data from SMILE has been additionally calibrated through a method known as alignment (Morrison, 2006). Through the alignment process some of the original survey attributes of each individual are tweaked slightly. For example, in order to meet an exogenous count of the ED-level distribution of labour force status by age and sex, a number of the LLI survey individuals chosen to represent that ED have their labour force status changed using a statistical imputation process. The other survey attributes aligned for this paper were income (county-level) and health (county-level). Alignment brings the simulated dataset into line with exogenous data not used as part of the original reweighting process, whilst maintaining a match with the original reweighting constraints. This introduces the required exogenous multivariate distributions within the simulated dataset.

Due to the way in which SMILE is created, each individual in the LLI will not be present more than once in each ED, but will reappear multiple times across the EDs of Ireland (because the 13800 individuals in the LLI have to represent the 3.2 million people of Ireland). Prior to alignment this potentially under-represents the variability of individuals between EDs. However, given the spatial resolution and number of attributes involved in the alignment process, few individuals from the original LII remain as identical replicas of each other in the final SMILE dataset. This reintroduces the potentially missing between-area variability of the real population. Thus, analysis on the post alignment dataset can use an individual level regression analysis, such as logistic regression models without worrying about survey clustering effects. Morrissey et al., (2013) provide a full discussion on the calibration process and subsequent validation of the health component of SMILE.

Of particular interest to this paper is that SMILE contains a health component with a self-reported depression variable for each individual in the Irish population (Morrissey et al., 2010). SMILE, however, does not contain a corresponding acute psychiatric usage variable for each individual. Statistically matching the data on rates of depression produced by SMILE (Morrissey et al., 2010) with the NPIRS data can provide a fully calibrated dataset on self-reported depression and admissions to acute psychiatric hospital data for depression. To achieve this statistical match, this paper proposes the use of propensity score matching (PSM).

3.3 Methodology

The use of PSM in healthcare research has increased over the last two decades particularly in the area of programme evaluation (Austin et al., 2005; Austin and Mamdani, 2005). However, recent work on data merging has seen an increase in the use of PSM to merge datasets with common variables (von Randow et al., 2012; Abello et al., 2008). PSM allows individuals admitted with depression from the NPIRS dataset to be matched to an individual reporting depression in the SMILE dataset based on the similarity of their demographic, socioeconomic and spatial characteristics. Of the 300,000 individuals in the SMILE dataset that reported suffering from depression in 2002, 5,113 (according to the NPIRS dataset) were admitted to a psychiatric facility. Thus 5,113 or approximately 2% of individuals with depression in the SMILE dataset need to be assigned an in-patient status.

3.4 Common Variables

To link two datasets using PSM the first step involves ensuring that that there are a number of common variables between the NPIRS and SMILE datasets. Table 1 presents the variables contained in the newly created SMILE admissions dataset using PSM. Examining Table 1 it is important to note that although marital status and employment status is included in the NPIRS dataset, the data records are incomplete. Thus, these two variables were not used in the PSM algorithm. In this paper, we match across four common variables, the depression variable (the treatment variable), age, gender and ED location. These were the only variables recorded for 100% of in-patients in the NPIRS dataset. Within the scope of these four variables, some testing around the choice of matching variables was conducted. Initially, the algorithm was run using only ED as a matching variable. It was found that the demographic and socio-economic profile of those assigned an in-patient admission status in the augmented SMILE dataset did not accurately reflect the profile of individuals that had been recorded in NPIRS. In other words, the ED-level distribution of age (as a continuous variable), gender and depression status did not match the ED-level age and sex distribution observed in the NPIRS dataset.

Table 1 Data contained in the SMILE Admissions Model

SMILE Admissions	SMILE	NPIRS
Depression	√	
Admission to an APH for depression		√
Sex	√	√
Age	√	√
Marital Status	√	√
ED Location	√	√
Distance to APH (via GIS)	√	
Household Income	√	
Education Level	√	
Employment Status	√	√
Occupation	√	
Medical Card Cover	√	
Urban/Rural Location	√	√
No. of Outpatient Facilities in region		√

No. of day Facilities in region		√

A second test involved using ED, depression status, sex, and single year of age. This achieved a match of 98% (4,935 out of the 5,113 individuals). That is, some of the EDs in SMILE contained insufficient individuals of a particular single year of age to precisely match the number of persons of that age, sex and ED recorded in NPIRS. In particular, it was elderly individuals (75 years plus) that proved the most difficult to match. Given the importance of outliers in microsimulation models (Birkin and Clarke, 2012) it was therefore decided to band the age variable into seven categories. This test achieved yielded an age-band match rate of 100%. Given the minimal loss of information on age distribution this entailed, it was therefore decided to use this approach for the remainder of the paper. As the NPIRS dataset and the SMILE dataset of depression are full populations, weighting considerations were not necessary.

A statistical loop was created so that the matching process was broken up on an administrative county by county basis. Thirty administrative counties were thus defined – the twenty five counties, Tipperary North and Tipperary South and the four administrative boroughs in Dublin. Each county was further broken up by gender, so that there were two county files, one female and one male. For example, only admissions data for females in Cavan in the NPIRS dataset were matched to females from Cavan with depression in the SMILE dataset. This spatially nested matching approach has two important outcomes. Firstly, partitioning the search by county increases the computational efficiency and effectiveness of the PSM process by decreasing the potential search space. Secondly, partitioning the search by county and using the geographical variable ED as an explicit matching variable significantly limits the introduction of spatial error into the augmented SMILE dataset. The underlying spatial variability of admissions in the augmented SMILE dataset was further ensured by excluding the 64% of EDs in SMILE that had no corresponding admissions in the NPIRS dataset. This ensured that only individuals within EDs that had an admission to an acute psychiatric hospital according to the NPIRS could be assigned an admissions status in SMILE.

3.5 Propensity Score Matching Algorithms

Figure 1 presents a schematic of the methodology employed. To begin the PSM process the 300,000 individuals with depression in the SMILE dataset were saved in their own file along with their depression status, age, sex and ED values. The 5,113 individuals that were admitted to an acute psychiatric facility in 2002 in the NPIRS, along with their depression diagnosis, age, sex and ED values, were saved in a second separate file. It is important to note that we only wish to assign an in-patient status to 5,113 of the 300,000 individuals with depression in the SMILE dataset. Given this requirement, and to ensure the correct multi-variate distribution of the non-matching variables for each individual, matching without replacement (von Randow et al., 2012; Abello et al., 2008) was employed. That is, each individual in the NPIRS dataset was only matched once to an individual from the SMILE dataset based on the similarity of the matching covariates.

Figure 1 Overview of Methodology

The next step of the PSM process as presented in Figure 1 involves appending the 5,113 individuals from the NPIRS dataset to the end of the 'depression only' SMILE dataset. PSM uses the 'propensity score' or the conditional probability of participation (treatment group, Y_1) to identify and match a counterfactual group of non-participants (outcome group, Y_0), given a set of observable covariates, X. Matching relies on the assumption of conditional independence. Individuals with similar propensities are matched and analysed pair-wise, so that given X, the outcome Y is conditionally independent of whether the individual received treatment. To match the NPIRS dataset to the SMILE dataset using PSM techniques, the NPIRS dataset is designated the treatment group (Y_1) and the SMILE dataset with the 300,000 individuals with depression is designated the outcome group (Y_0). The final step of the PSM technique as presented in Figure 1 is to assign treatment probabilities to each individual within the SMILE dataset. To assign treatment probabilities, a logistic or probit model is used in conjunction with a matching algorithm.

Several PSM methods exist including nearest neighbour, stratification, radius, kernel and local linear regression matching algorithms (von Randow et al., 2012; Abello et al., 2002; Jesmin et al., 2012). While there is no clear rule for determining which algorithm to use pre-estimation, using post estimation results it is possible to examine which algorithm best satisfies the balancing property. This means that observations with the same propensity score must have the same distribution of observable covariates, independent of treatment status. In this paper, three algorithms, nearest neighbour (NN), kernel density, uniform (KD) and local linear regression (LLR) were tested. Nearest neighbour matching is one of the most straightforward matching procedures. An individual from the comparison group (SMILE) is chosen as a match for a treated individual (in-patient, NPIRS) in terms of the closest propensity score (or the case most similar in terms of observed characteristics).

Kernel and local-linear regression matching are non-parametric matching estimators that compare the outcome of each treated person to a weighted average of the outcomes of all the untreated persons, with the highest weight being placed on those with scores closest to the treated individual. One major advantage of these approaches is the lower variance, which is achieved because more information is used. A drawback of non-parametric methods is that some of the observations used may be poor matches.

3.6 Algorithm Choice and Validation

Validation of the model outputs was conducted via a series of balance tests (Jesmin et al., 2012). Balance, with regard to PSM, is defined as the similarity between the multivariate empirical distributions of the covariates in the treated and control groups (Rosenbuam and Rubin, 1984). To test for balance in this paper, pseudo R^2, p-score and χ^2 test statistic were employed. Examining algorithm performance in terms of the reported pseudo R^2, p-score, and χ^2 all three algorithms demonstrated good covariate balance between the original NPIRS dataset and the newly created, admissions augmented SMILE dataset. However, the kernel algorithm slightly out-performed both the NN and LLR algorithm in that the values of the pseudo R^2, p-score and χ^2 test statistics, after matching, were lower than the values before matching. This confirms 'good balancing' since post matching there should be no systematic difference in the distribution of covariates between the two groups. Thus, the kernel-matching algorithm was chosen to perform the overall match.

To further test the robustness of the kernel-matching algorithm, Table 2 provides a comparison between matched individuals in the NPIRS and SMILE datasets. Every matched individual in SMILE had the same depression status, gender, age-band as their matched counter-part in the NPIRS dataset. For a small number of EDs within certain counties, no individual in the SMILE dataset matched the required age-band, gender, and depression profile. In consequence, the matching algorithm chose the next most similar individual, who had the required age, gender and depression profile, but was located in a neighbouring ED. Given that this only occurred in a small number of cases (1% of individuals) and given that the match was performed on a county basis with ED numbers spatially ordered in a sequential manner (i.e. ED 32001 is beside ED 34002), the impact on the access variables such as road distance to acute psychiatric hospital and availability of regional facilities is minimal. The results presented in Table 2 further confirm no obvious rural-urban bias in the performance of ED matching. Ninety-eight per cent of individuals that lived in a 'city' ED in the NPIRS dataset were assigned a 'city' ED in the augmented SMILE dataset; and 96% of individuals for rural EDs.

The NPIRS distribution was also successfully captured for those variables not used as part of the matching process. For example, 99.5% of the individuals matched between NPIRS and SMILE via PSM have the same single year of age. Similarly, the in work and unemployed, these variables have a matching rate of 88% and 87%, respectively. The final column of Table 2 identifies the minimum and maximum ED-level match rates. This additional analysis confirms that the achieved ED-level match rates for both the matched and non-matched variables vary little spatially.

Table 2 Comparison tabulation of matched and unmatched variable between the NPIRS and newly linked datasets

Variables	Original NPIRS	Matched (Kernel)	Match Rate	Range across EDs
Matching variables				
Depressed	100%	100%	100%	
Female	58%	58%	100%	
Age band				
14-24 years old	10%	10%	100%	
25-34 years old	15%	15%	100%	
35-44 years old	20%	20%	100%	
45-54 years old	20%	20%	100%	
55-64 years old	16%	16%	100%	
65-74 years old	12%	12%	100%	
75 plus years old	7%	7%	100%	

ED Location All City Rural	 27% 56%	 28% 55%	 99% 96% 98%	 97-99% 95-97% 96-98%
Non-matching variables				
Age (single year)	47.8 years old	48 years old	99.5%	99.5%
Labour Force Status In-work Unemployed	 26% 23%	 23% 20%	 88% 87%	 86-89% 85-89%

However, it is not good enough to merely adequately capture in the augmented SMILE dataset the observed univariate distributions in NPIRS. In order to subsequently explore in full the demographic, socio-economic and geographical factors underpinning in-patient admission it is also necessary for SMILE to capture as accurately as possible the interactions between the outcome (in-patient admission) and its potential underpinning factors.

Data are lacking to validate the full set of possible interactions of interest. However, it is at least possible to compare the traits of persons flagged as in-patients admitted for depressive illnesses in NPIRS and the augmented SMILE. With regard to the multivariate distribution for two matching variables, gender by urban/rural residence, 29%, 55% and 15% of females came from urban, rural and town areas (respectively) in the newly augmented SMILE dataset, whilst 27%, 57% and 18% of females came from urban, rural and town areas in the original NPIRS Dataset. With regard to males, 27%, 57% and 15% of males came from urban, rural and town areas in the newly augmented SMILE dataset, whilst 28%, 54% and 18% of males came from urban, rural and town areas in the original NPIRS Dataset. Similarly, for the non-matched variable unemployment, the SMILE dataset reported that 19% and 28% of depressive psychiatric in-patients were unemployed females and males, whilst the original NPIRS recorded 21% and 28% unemployed females and males, respectively. These cross-tabulations suggest that the multivariate distributions between key variables of interest were successfully replicated during the matching process.

Table 3 Cross-tabulations (percentage and counts) to validate multivariate distributions between the newly augmented SMILE dataset and NPIRS

Matched variables							Unmatched variable	
	Urban		Rural		Town		Unemployed	
Gender	SMILE	NPIRS	SMILE	NPIRS	SMILE	NPIRS	SMILE	NPIRS
Female	27% (841)	29% (904)	57% (1776)	55% (1714)	15% (467)	15% (467)	19% (592)	21% (654)
Male	28% (559)	25% (499)	57% (1138)	54% (1078)	18% (359)	18% (359)	28% (559)	28% (559)

Finally, the Kullback-Leibler divergence measure (K–L divergence) for the discrete case (von Randow et al., 2012) was used to formally compare the distribution of the characteristics among individuals who were allocated an in-patient status (SMILE) with individuals that had been in-patients (NPIRS). Using the K-L divergence statistic, it was found that the distribution and cumulative percentage of K-L divergences among in-patients in both datasets were identical (0.00). The sensitivity analysis and distributional analysis indicate that the newly matched SMILE in-patient variable is representative of the in-patient population reported in the NPIRS dataset.

In the next section of this paper, this augmented SMILE dataset is used to model in-patient admission given age-band, sex, work status, income, medical card status, education and a range of location-related measures. We have shown above that the relationship between in-patient status and age-band, location and sex has been well captured through the PSM process. This begs the question of how well the original SMILE dataset has captured the full multivariate and spatial distribution of these explanatory factors. All of these explanatory factors were used in either the reweighting or alignment stages of the SMILE dataset creation process. As Morrissey et al., (2013) and Morrissey and O'Donoghue (2011) demonstrate the resulting SMILE dataset adequately captures the spatial multivariate distribution required. In particular these evaluations threw up no evidence of systematic spatial bias. The fact that the PSM matching rates reported in this paper are so high perhaps provides some further reassurance about the quality of the SMILE dataset, indicating that the correct number of individuals by age, sex and depression status were available in the SMILE model.

4. Analysis

4.1 Analysis of the Logistic Model

Using the variables presented in Table 4, the relationship between admissions to acute psychiatric facilities and individual level locational, demographic and socio-economic characteristics was examined by running a logistic model on the newly matched data. It is important to note that the sample size of the logistic regression is 300,000 individuals. This represents the number of individuals identified as having depression within the general Irish population. This allows us to specifically identify how the characteristics identified by the literature as being possible important predictors covary with the likelihood of being admitted to an acute psychiatric hospital for depression. Table 4 presents the results of the logistic model. The calculated deviance statistic is 98.3%. Using Hosmer and Lemeshow goodness of fit statistic, the overall model was found to be statistically significant (Hosmer-Lemeshow $\chi^2 = 95.6$ and $p = .00$). This means that including the variables outlined in Table 4 fits the matched data statistically significantly better than a model without explanatory variable (only the constant). Wald tests, represented as confidence intervals within Table 4, were used to test the statistical significance of each variable at the 95% level. Odds ratios for each of the explanatory variables are also included in Table 4, however, the reported standard errors and confidence intervals are for the log odds co-efficient. First order-interaction interactions were also examined but none of them were found to be significant or improved the model's fit according to the Hosmer and Lemeshow goodness of fit statistic.

Examining the results presented in Table 4, the constant in the logistic regression model shows the odds of an individual with depression being admitted to an acute psychiatric hospital is less probable for the reference group (females, residing in towns, aged 14 to 24 years old, with only primary education, not employed, who do not have access to free medical care in Ireland (medical card holder)). From Table 4 it can be seen that females (-0.149), higher age categories (relative to the lowest age band, 14 to 24 years old) and higher levels of education, except for university level (relative to primary school education only) have a positive relationship with admission to an acute psychiatric hospital.

Table 4 Logistic Model of Admission to an APH for individuals with Depression

Admission to a Psychiatric Hospital	Co-efficient	Standard Error	95% Confident Intervals		Odds Ratios
Rural residence (reference town residence)	-0.293	0.042	-0.376	-0.211	0.77
City residence (reference town residence)	-0.336	0.066	-0.465	-0.206	0.63
Distance to a APH	-0.002	0.001	-0.004	0	1.00
Number of Day facilities in Region	-0.032	0.008	-0.046	-0.017	0.98
Number of Outpatient facilities in Region	0.006	0.002	0.001	0.011	1.00
Age 24-34 (ref. Age 14-24)	0.677	0.061	0.558	0.797	1.06
Age 35-44 (ref. Age 14-24)	1.149	0.06	1.032	1.266	1.12
Age 45-54 (ref. Age 14-24)	0.843	0.059	0.727	0.959	1.53
Age 55-64 (ref. Age 14-24)	0.538	0.062	0.417	0.659	0.99
Age 65-74 (ref. Age 14-24)	0.425	0.069	0.291	0.559	0.93
Age 75 plus (ref. Age 14-24)	0.171	0.078	0.019	0.324	0.88
Male	-0.149	0.031	-0.21	-0.088	0.73
Work Status (In work - 1)	-0.138	0.041	-0.219	-0.058	0.82
Household weekly income €24,000-€34,000 (ref. < €24,000)	0.145	0.056	0.035	0.254	1.16
Household weekly income €35,000-€44,000 (ref. < €24,000)	0.149	0.039	0.072	0.225	1.61
Household weekly income €45,000-€60,000 (ref. < €24,000)	-0.309	0.06	-0.427	-0.191	1.00
Household weekly income €60,000 plus (ref. < €24,000)	0.13	0.136	-0.138	0.397	1.30
Medical Card Status (1 - has medical card)	0.118	0.037	0.046	0.191	1.07
Lower Secondary Education (relative to primary only)	0.493	0.046	0.403	0.583	1.00
High Secondary Education (relative to primary only)	0.468	0.052	0.367	0.57	1.35
Diploma (relative to primary only)	0.222	0.061	0.102	0.342	1.14
University Degree or higher (relative to primary only)	0.078	0.088	-0.095	0.251	0.93

Constant	-4.293	0.164	-4.614	-3.971	

This analysis uses three variables as proxies for socio-economic status including, equivalised household income, employment status and medical card status. Examining equivalised household income, each band has a strong positive relationship relative to the lowest income band (less than €24,000), except for the income band €45,000-€60,000. Individuals with depression in this income band have a negative relationship with admissions to an acute psychiatric hospital. Entitlement to free health care in Ireland (medical card holder) may serve as a proxy for disadvantage in Ireland (Kelleher et al., 2003). The logistic model presented in Table 4 found a significant positive relationship between individuals with a medical card and admissions to an acute psychiatric hospital (0.118). This reinforces research by O'Keane et al., (2004) that the spatial distribution of medical cardholders is a key determinant of the use of public acute psychiatric hospitals in Ireland.

Examining employment status, the logistic model found that individuals with depression not in work (-0.135) have a significant relationship with admissions to acute psychiatric facilities. However, it is important to note that whether an individual is not in work because of their illness or lack of work has resulted in depression, cannot be inferred from these findings. The logistic model found that individuals with depression from rural (-0.293) and city (-0.336) residences have a negative covariation with admissions to acute psychiatric facilities relative to those living in towns. Thus, comparable with previous international research (Van de Velde et al., 2010), this paper found that, being female, single and older had a significant positive relationship with admissions to an acute psychiatric facility for depression. However, socioeconomic status had an amibiguous relationship with admissions (Lorant et al., 2003)

With regard to the provision of psychiatric services, three variables, road distance to an acute psychiatric facility, the number of day facilities in a region and the number of outpatient facilities in a region, were included in the model. In relation to the community based provision of psychiatric services, it was found that whilst a higher number of day facilities in a region covaried negatively with admissions to acute psychiatric facilities, a higher number of outpatient services were positively associated with admissions. The negative relationship between day facilities and admissions is to be expected. However, it might also be expected that this would be true for outpatient services. Two contrasting reasons may be presented for this finding. Firstly, individuals with depression in areas with better access to outpatient services (a higher number of facilities) may be more easily referred to an acute psychiatric facility, whilst individuals with lower access are 'under-serviced' and their needs are left unmet. Previous research in Ireland noted that a significant number of admissions to psychiatric services via Accident and Emergency departments are frequent attendees (Okorie, et al., 2011).

The second explanation may be that consultants in regions with more outpatient facilities are over-referring to acute psychiatric facilities regardless of need. This result is interesting and data outside of the SMILE admissions model is required to investigate the relationship further. Examining access in terms of road network distance, it was found that an increase in distance has a significant negative covariation with admissions. That is, individuals with depression closer to an acute psychiatric facility have a stronger association with admissions to such a facility. Whilst it is not possible to say with certainty that access to acute psychiatric facilities has an impact on whether an individual with depression seeks acute psychiatric care, it can be inferred that there is a correlation between access and admissions.

4.2 Sensitivity Analysis

The results above are based upon the SMILE dastaset augmented with NPIRS admission status data using four matching variables: ED, age-band, sex and depression status. To test the sensitivity of these findings to the variables used for matching, two alternative versions of the augmented SMILE dataset were created, using different combinations of matching variables. Equivalent logistic regression models were then fitted to these variant SMILE datasets. Running the logistic model using ED, *single year* of age, sex and depression status led to no difference of note compared to the results presented in Table 4. In contrast, running the logistic model on SMILE data matched using only ED resulted significantly different results. The access variable became statistically insignificant. Furthermore the relationship between age and marital status (single or married) and admissions became insignificant, and with regard to age, negative. This is in contrast to the significant positive covariation of age and negative covariation between marital status admissions to acute psychiatric facilities for depression well documented in the literature (Fone et al., 2006; Peterson et al., 2009; Van de Velde et al., 2010). This perhaps confirms the earlier conclusion that using only ED as a matching variable results in a poor replication of the demographic and socioeconomic distribution for individuals with an admission to an acute psychiatric facility within the SMILE dataset.

5. Discussion

The aim of this paper was to examine the difference in individual and locational characteristics between individuals with depression and individuals with depression who are admitted to an acute psychiatric facility. Using newly matched data and a logistic model, this paper wished to examine if ease of access was an important factor in determining if individuals with depression are admitted to an acute psychiatric hospital. Using road network distance and controlling for a number of demographic and socio-economic factors it was found that individuals with depression that live closer to an acute psychiatric facility are more likely to be admitted to such a facility. Previous research, particularly on the role of 'urbanicity' on mental health (Peen et al., 2009) has indicated that increased admissions related to better access may be due to three reasons. First, individuals with better access may use services on a more frequent basis, whilst individuals with poorer access, although in need of care, will remain 'under-serviced' (McLafferty, 2003). This would indicate a clear role for public policy in ensuring better access to health facilities (McLafferty, 2003). Second, the significant role of access may be due to differing admission practices to psychiatric in-patient services rather than true morbidity differences between areas (Tedstone-Doherty et al., 2007). Practitioners in areas with good access may admit individuals at a higher rate than practitioners where access is poor. Third, it may be that individuals most in need of care, individuals with depression 'drift' towards areas with higher health service provision to ensure ease of access. This is known as the 'social drift hypothesis' (Peen et al., 2009; Kilbride et al., 2010).

5.1 Using Matched Data for Statistical Analysis

In interpreting the model results, it should be noted that there are practical assumptions and decisions made in the PSM process (Von Radon et al., 2011). These decisions tend to be driven by what is available in the data at hand while also trying to maintain rigor. In this paper, there were a limited number of variables within the NPIRS dataset that could be used with certainty. This limited the common variables between the two data sets to which the matching process could be employed. It is also important to note that matching algorithms rely on the assumption of conditional independence (Morrissey and O'Donoghue, 2012; von Radon et al., 2011); all variation in admissions to an acute psychiatric hospital can be explained by the matching variables. To the extent that this does not hold – for example if the interaction between educational status, the matching variables and the propensity for a person suffering from depression to be admitted to an acute psychiatric hospital varies significantly across space - the model results will have a tendency to understate area differences.

Directly linked to the assumption of condition independence is the use of simulated data and predictive statistical models together. Predictive models are based on the assumption that certain variables in the equation of interest are missing. Thus, all the results provided by these models are suggestive and depend on the choice of variables included in the model. Models estimated using simulated data therefore contain the error associated with statistical matching and the error associated with predictive models. However, Hynes et al. (2010) have noted that the level of error from models estimated from simulated data can be reduced if the researcher ensures that the statistical matching variables are believed *a priori* to be useful in predicting the dependent variable in the model in question, or that the matching variables are determinants of the explanatory variables in the model. The matching variables used included age and gender, which are known to be the key determinants of depression (Morrissey et al., 2010) and admissions to acute psychiatric facility for depression (Morrissey et al., 2013). The use of a geographical variable, ED as an explicit matching variable significantly limits the introduction of spatial error in the augmented SMILE dataset. In particular, persons from a given ED will necessarily have been assigned the correct values for the ecological variables in the model relating to urbanicity, road distance to acute psychiatric hospital and availability of regional facilities. However, as with all statistical modelling approaches the evidence presented here should be viewed as suggestive and not inferential.

Limitations aside, to disentangle the effect of access to acute psychiatric hospital on admission rates for individuals with depression requires further analysis outside the framework of data provided by the matched SMILE/NPIRS dataset. These results provide the first evidence that access to an acute psychiatric facility in Ireland is correlated with admissions to these facilities for individuals with depression. Although this finding is country-specific, it was only through the use of PSM to link a spatial microsimulation with an administrative dataset that such an analysis was achievable. In recent years, a number of statistical agencies (e.g. the Australian Bureau of Statistics, ABS) have used sample surveys to enable richer data to be obtained at a cheaper cost, and with less total respondent burden, than using a census of the whole population (Tanton et al., 2011). However, one of the problems with official sample surveys is that samples are designed to provide estimates regions or government authorities, but not for areas below this level. Small area estimation techniques, such as spatial microsimulation, overcome this limitation to a great extent and may provide an alternative to census based small area population estimates. Illustrating the use of PSM in matching administrative data on acute psychiatric service usage to SMILE provides a framework for the addition of administrative data to large scale surveys in a computationally efficient and robust manner. Thus, this paper demonstrates that through PSM existing data can be updated for further analysis or modelling that otherwise would not have been possible, potentially alleviating the need for further expensive data collection.

References

Abello A, Lymer S, Brown L, Harding A, & Phillips B, 2008, "Enhancing the Australian National Health Survey for use in a microsimulation model of pharmaceutical drug usage and cost". *Journal of Artificial Societies and Social Simulation* **11**(2).

Almog M, Curtis S, Copeland A, Congdon P, 2004, "Geographical variation in acute psychiatric admissions within New York City 1990–2000: growing inequalities in service use?" *Social Science & Medicine* ***59(2)*** 361–376.

Austin PC, Mamdani MM, Stukel TA, Anderson GM, Tu JV 2005, "The use of the propensity score for estimating treatment effects: Administrative versus clinical data" *Statistics in Medicine*, **24**(10) 1563‐1578.

Austin PC, Mamdani MM 2006, "A comparison of propensity score methods: A case‐study estimating the effectiveness of post‐AMI statin use" *Statistics in Medicine* **25**(12) 2084‐2106.

Birkin M, Clarke G.P, 2012, "The enhancement of spatial microsimulation models using geodemographics" *Annals of Regional Science*, **49**, 515–532.

Congdon, P, 2002, "A Model for Mental Health Needs and Resourcing in Small Geographic Areas: A Multivariate Spatial Perspective" *Geographical Analysis, 34(2)*, 168-186.

Curtis SE, Copeland A, Fagg J, Congdon P, Almog M, Fitzpatrick J, 2006, "The ecological relationship between deprivation, social isolation and rates of hospital admission for acute psychiatric care; a comparison of London and New York City" *Health and Place* **12**(1) 19-37.

Curtis S, 2007, "Socioeconomic status and geographies of psychiatric inpatient service use. Places, provision, power and wellbeing" *Epidemiologia e Psichiatria Sociale* **16** 10-15.

Daly A, Walsh D, 2012, *Activities of Irish Psychiatric Units and Hospitals 2011*, Health Research Board Statistics Series 18, Health Research Board, Dublin.

Daly A, Walsh D, 2011, *Irish Psychiatric Units and Hospitals Census 2010*, Health Research Board Statistics Series 12, Health Research Board, Dublin.

Department of Health and Children, 2006, "A Vision for Change" Government of Ireland, Dublin, Ireland.

Department of Health, 1984, "The Psychiatric Services – Planning for the Future" The Stationary Office, Dublin, Ireland.

Fone D, Dunstan F, Williams G, Lloyd K, Palmer S, 2007, "Places, people and mental health: A multilevel analysis of economic inactivity" *Social Science & Medicine*, **64** 633–645.

Fortney J, Rost K, Zhang M, Warren J, 1999, "The impact of geographic accessibility on the intensity and quality of depression treatment *Medical Care **37**(9)* 884–93.

Jesmin S, Thind A, Sarma S, 2012, "Does team-based primary health care improve patients' perception of outcomes? Evidence from the 2007–08 Canadian Survey of Experiences with Primary Health" *Health Policy* **105**(1) 71-83.

Joseph A, Phillips D, 1984, *Accessibility and utilization: Geographical perspectives on health care delivery*. London, UK: Harper & Row.

Kelleher M, O'Brien S, 2001, "An audit of admissions to an acute psychiatric unit', *Irish Journal of Psychological Medicine*, **18**(4) 132-136.

Kelleher C, Timoney A, Friel S, 2003, "Socio-demographic predictors of self-rated health in the Republic of Ireland: findings from the National Survey of Lifestyle, Attitudes and Nutrition, SLAN" *Social Science and Medicine*, **57** 477-86.

Kirkbride J, Coid J, Fearon P, Morgan C, Dazzan P, Yang M, Morgan K, Lloyd T, Harrison GL, Murray RM, Jones PB, 2010, "Translating the epidemiology of psychosis into public mental health: evidence, challenges and future prospects" *Journal of Public Mental Health* **9**(2) 4-14.

Koppel S, and Mcguffin P, 1999, "Socioeconomic factors that predict psychiatric admissions at a local level" *Psychological Medicine* **29** 1235-1241.

Lorant V, Deliege D, Eaton W, Robert P, Phillippot P, Ansseau M 2003, “Socioeconomic Inequalities in Depression: A Meta-Analysis” *American Journal of Epidemiology* **157**(2) 98-112.

Luo W, Wang F, 2003, "Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region" *Environment and Planning B: Planning and Design* **30**(6) 865 – 884

McCarthy JF, Blow FC, Valenstein M, Fischer EP, Owen RR, Barry KL, Hudson TJ, Ignacio RV, 2007, “Veterans Affairs Health System and mental health treatment retention among patients with serious mental illness: Evaluating accessibility and availability barriers” *Health Services Research* **42** 1042-1060.

McLafferty S, 2003, “GIS and Health Care”, *Annual Review Public Health* **24** 25–42.

Morrissey K, Clarke G.P., Ballas D, Hynes S, O’Donoghue C, 2008, “Examining access to GP services in rural Ireland using microsimulation analysis” *Area* **40** 354-364.

Morrissey K, Hynes S, Clarke G.P., O'Donoghue C, 2010, “Examining the factors associated with depression at the small area level in Ireland using spatial microsimulation techniques” *Irish Geography*, **43**(1) 1- 22.

Morrissey K, O'Donoghue C, 2013, “Validation Issues and the Spatial Pattern of Household Income”, In O'Donoghue, C., Ballas, D., Clarke, G.P., Hynes, S., Morrissey, K. (Eds.), *Spatial Microsimulation for Rural Policy Analysis*, London, Springer.

Morrissey K, Daly A, Clarke G.P., Ballas D, O’Donoghue C, 2012, “A Rural/ Urban Comparison of Psychiatric Inpatient Admissions in Ireland” *Journal of Public Mental Health* **11**(4) 209-213.

Neng Wan F, Benjamin Zhan, Bin Zou, Edwin Chow, 2012, “A relative spatial access assessment approach for analyzing potential spatial access to colorectal cancer services in Texas” *Applied Geography 32(2)* 291-299.

O’Keane V, Jeffers A, Moloney E, 2004, “Irish Psychiatric Association survey of psychiatric services in Ireland” *Psychiatric Bulletin* **28**, 364–367.

Peen J, Schoevers R.A, Beekman A.T, Dekker J, 2010, “The current status of urban-rural differences in psychiatric disorders” *Acta Psychiatric Scandinavi* **121** 84–93.

Peterson LE, Alexander C, Tsai S, Petterson D, Litaker G, 2009, “Rural-urban comparison of contextual associations with self-reported mental health status” *Health and Place* **15** 125-132.

Pfeiffer P, Glass J, Austin K, Valenstein M, McCarthy J, Zivin K, 2011, “Impact of Distance and Facility of Initial Diagnosis on Depression Treatment” *Health Services Research* **46** 768–786

Rosenbaum PR, Rubin DB, 1984, “Reducing bias in observational studies using subclassification on the propensity score” *Journal of American Statistics Association* **79** 516–24.

Rushton G, and M West, 1999, "Women with localized breast cancer selecting mastectomy treatment, Iowa, 1991-1996" *Public Health Reports* **114**(4) 370-1.

Smith DM, Pearce JR, Harland K, 2011, “Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand” *Health and Place* **17**(2) 618-624.

Tedstone-Doherty F, Walsh D, Moran R, Bannon, 2007, *High Support Community Residence Census 2006*, Health Research Statistics Series 1, Mental Health Research Unit, Health Research Board, Dublin.

Tedstone-Doherty D and Moran R, 2009, Mental health and associated health service use on the island of Ireland. HRB Research Series 7. Dublin: Health Research Board.

Von Randow M, Davis P, Lay-Yee R, Pearson J, 2012, “Data Matching to Allocate Doctors to Patients in a Microsimulation Model of the Primary Care Process in New Zealand” *Social Science Computer Review* **30**(3) 358-368.

Van de Velde S, Bracke P, Levecque K, 2010, “Gender differences in depression in 23 European countries. Cross-national variation in the gender gap in depression” *Social Science & Medicine* **71**(2) 305-313.