Research Article

Developing a Dynamic Microsimulation Model of the Australian Health System: A Means to Explore Impacts of Obesity over the Next 50 Years

Sharyn Lymer and Laurie Brown

National Centre for Social and Economic Modelling, University of Canberra, Canberra, ACT 2601, Australia

Correspondence should be addressed to Sharyn Lymer, sharyn.lymer@natsem.canberra.edu.au

Health of the ageing population has the potential to place considerable pressure on future government spending. Further, the impacts of the obesity epidemic have the potential to place additional pressure on government health budgets. In response to such fiscal concerns in Australia, a dynamic microsimulation model, APPSIM, has been developed at the National Centre for Social and Economic Modelling (NATSEM). The health module was developed to allow consideration of health behaviours within the context of an ageing population and the resultant health profile of the population. Also included in the modelling is the associated use of health services and their costs. All health variables used were imputed onto the 2001 basefile derived from the 1 percent unit record file of the 2001 Australian census. Transition equations of these variables were estimated to allow projections over time. In this paper, the model has been used to look at the impacts of obesity on the Australian population health profile and associated health expenditure. In the scenario, removal of obesity from the population leads to a simulated population with a better health profile but showed only marginal changes in relative health expenditure.

1. Introduction

It is well known that the Australian population is ageing and that across all age groups there is rising levels of obesity. In 1971, 8 percent of the Australian population was aged 65 years and over: by 2010, this had increased to almost 14 percent [1]. Official projections indicate that by 2050 some 23 percent of the Australian population will be aged 65 years and over [2]. An ageing population places increased pressure on government spending through increased demand for health care, aged care, and pensions. Health care spending has been steadily growing, from $Au 42 billion in 1996-1997 to $Au 103 billion in 2006-2007 [3]. Projections estimate continued rises in health expenditure from 3.7 percent of GDP in 2009/10 to 7.0 percent of GDP in 2046/47 [2].

Beyond the number or proportion of the aged population, the impacts on future health expenditure will be moderated by the health experience of the aged population. Possibilities of morbidity compression [4], expansion [5], dynamic equilibrium [6], or some cyclic effect between compression and expansion of morbidity [7] will impact the possible demand for health services. The relationship between health and longevity may be effected by the severity of disease not being as great due to slower progression of disease [8]. Further, issues such as new technology, medications, and changes in medical practice will mediate the effects of ageing on health expenditure [2, 9, 10].

The health of the population is also mediated by behavioural risk factors, such as obesity. Obesity has been increasing in Australia, going from 11 percent in 1995 [11] to 25 percent in 2007/08 [12]. Severe obesity is associated with chronic disease such as cardiovascular disease, hypertension, diabetes, some cancers, and arthritis. The relationship between obesity and mortality is complex, with studies having varied findings regarding the impacts of obesity on mortality. Direct impacts of obesity on mortality have not always been substantiated [13]. However, other studies have found direct impacts between obesity and mortality [14–17]

though this may have been influenced by variables controlled for in the analysis. Indirect impacts between obesity and mortality via links between obesity and chronic disease have been found in a range of studies [17–20]. Younger cohorts have increased weight at younger ages and over a shorter time frame [21–25]. Potentially the baby boomers will be a fatter, less healthy cohort than seen at any time previously. Whilst in Australia public policy debate surrounding the obesity epidemic amongst children has occurred [26] and adults generally [27, 28], less consideration of public policy to address and cope with obesity amongst the older population has occurred.

This paper will describe the construction of a health module for an Australian dynamic microsimulation model and explore the impacts of obesity on government spending. This modelling allows the consideration of obesity, health service usage, and expenditure using a bottom-up approach. The modelling allows long-term projections of obesity based on individual characteristics, tying the impacts of obesity to health status and consequent health service usage and expenditure.

2. Method

The key of this paper is the description of a general population microsimulation model and in particular the health module that has been developed to date.

2.1. Microsimulation. Microsimulation is a simulation technique which uses microlevel units as the unit of analysis [29–31]. It considers the decision-making microunit, usually an individual or family, within a large complex system. The key uses for microsimulation have been analysis of government policy to consider distributional impacts and allowing the determination of the winners and losers as well as how much does the individual gain or lose by the policy change. In particular, this model is a dynamic model. Dynamic microsimulation moves individuals through time updating each attribute for individual for each time interval, based on probabilities determined from appropriate data sources. The records (which represent the individuals) in the model are continuously updated, such that the records in the former time period become the starting point in the latter one. The changes over time are determined by stochastic Monte Carlo procedures based on the defined transition probabilities.

Dynamic microsimulation of health has not been attempted for Australia in a large general-purpose dynamic microsimulation model. Typically, general-purpose dynamic microsimulation models treat health in a simplistic manner such as SAGE [32] where health is a binary indicator variable. Alternately the models are built to explore specific aspects of health conditions and associated systems. These models do not consider the wider socioeconomic context (e.g., stroke prediction in a population with carotid artery stenosis [33] and colorectal cancer simulated population model for incidence and natural history [34]). Two notable exceptions are POHEM developed by Statistics Canada [35] and SESIM [36] developed in Sweden. However, these two models take very different approaches to the modelling process. POHEM is based on a synthetic basefile, and the equations operate in continuous time. SESIM is based on an administrative database and operates in discrete time (annual assessment of individual states). POHEM has been used primarily to consider cost-benefit aspects of treatment for specific diseases such as cancer [37–39], cardiovascular disease [40], and osteoarthritis [41]. SESIM in its 2007 update included for the first time a health status indicator. In its latest upgrade, health service and aged care use have been modelled [42]. The health modeling for inclusion in APPSIM is most similar to that reported in SESIM.

2.2. APPSIM. The dynamic microsimulation model, "Australian Population and Policy Simulation Model" (APPSIM) [43], provided the basis for the development of the health module that is to be described. APPSIM is a closed, population-based, dynamic microsimulation model that operates in discrete time. The model provides snapshots of the Australian population characteristics and government programs as at 30 June each year from 2002 to 2050. As a basefile, it used the publically released confidentialised 1% unit record file from the Australian Bureau of Statistics (ABS) 2001 Census, which provides approximately 180,000 individuals' records that can be included in the simulation.

The model functions in one-year time units, using discrete time functions. APPSIM is made up of a series of modules that are processed sequentially. Starting at time t, it moves through the modules sequentially starting with disability, demographics, household formation, education, labour force, earnings, housing, other income, household assets and debt, social security and taxation, health, and aged care (details of the development of the modelling within each module are available at the NATSEM website http://www) .canberra.edu.au/centres/natsem/publications/).

APPSIM is structured such that it has an initial starting population, a simulation cycle, and output. The output from the APPSIM model provides both high-level summary statistics and databases similar to survey data files. Consequently, there is the capacity to do three levels of analysis from the simulation output: time series, cross-sectional analysis looking at distribution impacts in certain years, and finally life-course analysis considering the impacts of changes to individuals over time periods.

The model has been programmed in c# and the parameters of the model are read via excel spreadsheets. For some outcomes in the model, such as births, unemployment, or educational retention rates, it is possible to use alignment so that the outcomes match ABS or treasury projections as a means of comparison. Alternately, alignment can be used to consider various counterfactuals.

2.3. Health Module. The primary focus of this paper is the health module that was developed for APPSIM. Due to the complexity of the health system, strategic simplification of the system was required to facilitate the modelling. The simplification was based on the aims of the model, as well

as the more practical considerations of human and financial resources and data availability.

The modelling is focused on adults, that is, persons aged 15 years and over, due to the lack of longitudinal data in Australia about children's health. The system included four health risk behaviours: inadequate physical activity, at-risk alcohol consumption, smoking status, and obesity. These interact with each other and also become input parameters for the modelling of health status. As well, socioeconomic factors such as education, income, and labour force status (which are modelled external to the health module) are input parameters to both the health risk behaviours and health status. In other modellings, health status becomes an input parameter to health service usage, which in turn is used in estimating health expenditure.

2.3.1. Data Source. The development of the health module involved integration of multiple data sources. The primary data source for the modelling within the health module was the HILDA survey [44]. It provided the data to model health risk behaviours, health status, and private health insurance. The HILDA survey is a complex sample survey that selected a panel of Australian households for followup. In the initial wave (i.e., 2001), approximately 14,000 individuals were interviewed. By 2008, there were just over 9,000 respondents. Due to the longitudinal nature of HILDA, it has allowed modelling of transition probabilities that account for an individual's state change over time. A major modelling restriction was the late inclusion of obesity within the survey (it has only been asked since the 2006 survey). Consequently, obesity and health status are modelled over three time points, whereas the other health behaviours are modelled over six or seven time points (with obesity excluded from the set of inputs).

For the modelling of health service usage and health expenditure, data is less rich, and consequently several data sources have been combined to achieve the modelling. A combination of administrative data (supplied by the Australian Department of Health and Ageing) and survey data (specifically the ABS National Health Survey) was used in the allocation of health service usage. Further, for modelling number of pharmaceutical scripts, information was also used from the NATSEM microsimulation model, MediSim [45]. For health expenditure on medical services and pharmaceutical scripts, administrative data informed the average costs per age-sex group of a service or script whilst for hospital admissions a combination of reported administrative data from the Australian Institute of Health and Welfare (AIHW) and information from the NATSEM hospital microsimulation model HospMod [46] was used to determine who had hospital visits, how many admissions they had, and for how long each hospital admission lasted.

2.3.2. Variable Definitions. For consistency, a set of definitions surrounding health risk behaviours and health were developed. The definitions where possible were based on recommendations and guidelines released within Australia for achieving good health (see Table 1). The health risk behaviours were all modelled as dichotomous variables. The health status indicator was based on a self-reported health question in HILDA. It is a 5-point indicator where the individual ranks their health between excellent and poor.

2.3.3. Baseline Imputation. The APPSIM basefile did not have individual health information, resulting in all the health risk behaviours and health status having to be initially imputed. A summary of the statistical methods and explanatory variables used for each outcome in the health module is presented in Table 2. All of the dichotomous variables used logistic regression to inform the probabilities of individual characteristics. The general equation for determining probabilities from a logistic regression is presented in Box 1. For health status, which had five levels, probabilities were calculated using generalised ordinal logistic regression [47]. The equations to determine the probabilities from this type of model are presented in Box 2. Explanatory parameters in all cases were a variety of socioeconomic variables. In addition, for health status, the health behaviours were also included as explanatory parameters within the system. For each of the health risk behaviours and health status, the statistical models were developed for males and females separately.

2.3.4. Transition Probabilities. To move the individuals through time, transition probabilities were determined using statistical models. The details of the type of model and explanatory parameters included in the model are presented in Table 3. The binary variables transition probabilities were calculated using pooled dynamic logistic regression models. The models were based on the HILDA data across several years being combined and treated records as being independent. The dynamic aspect is the inclusion of lagged outcome variables on the right-hand side of the model. In most cases, a selection of socioeconomic variables was included as explanatory variables, as well as the lagged dependant variable. The probability equations for this model are the same as presented in Box 1 with the inclusion of the lagged dependant variable. The transition probabilities for health status were estimated using a dynamic generalised ordinal logistic regression model with socioeconomic explanatory variables as well as health risk factors and health status in the previous time period. For this model, the probability equations have the same form as presented in Box 2. The detailed model specification including coefficient values is described in [49]. Separate models were estimated for men and women. As with all simulation modelling, there is a tension between the level of complexity of the statistical model and the accuracy of fit to the data whilst maintaining the ability to apply the statistical model within a microsimulation model.

2.4. Counterfactuals. To demonstrate the potential capacity of the modelling described in the previous sections, a baseline scenario and counterfactuals changing obesity levels within the population were run. The primary counterfactual looked at the impact of removing obesity from the population. An alternative with all the population being obese also was run to allow comparative analysis.

TABLE 1: Models for baseline imputation for health module.

Outcome	Definition
Inadequate physical activity	Less than three sessions of at least 30 minutes moderate activity [60]
At-risk alcohol consumption	Over 18 years: more than two standard drinks in a day Under 18 years: any alcohol consumption [61]
Smoker	Any reported volume of smoking
Obese	Body mass index greater than 30 [62]
Health status	Self-reported health ranked between excellent and poor on a 5-point scale
Private health insurance	Any expenditure on private health insurance coverage in the last year
Health service usage	Medical: number of MBS[1] services per person per year Pharmaceutical: number of PBS[2] scripts per persons per year Hospital: number of hospital admissions per person per year, and number of bed days per hospital admission
Health service costs	Medical: costs per MBS service Pharmaceutical: cost per PBS script Hospital: cost per hospital bed day

[1]MBS refers to the medical benefits scheme which provides government funding for medical practitioner services, optometry, diagnostic imaging, and pathology tests.
[2]PBS refers to the pharmaceutical benefits scheme which provides government subsidy to achieve "affordable access to necessary and cost-effective medicines" [7].

TABLE 2: Models for baseline imputation for health module.

Outcome	Statistical model	Explanatory variable
Inadequate physical activity	Logistic regression	Age, marital status, education, labour force status, kids in household, adults in household, and household income
Smoker	Logistic regression	Age, marital status, education, labour force status, kids in household, adults in household, and household income
At-risk alcohol consumption	Logistic regression	Age, marital status, education, labour force status, kids in household, adults in household, household income, and smoking status
Obese	Logistic regression	Age, marital status, education, labour force status, kids in household, adults in household, household income, and inadequate physical activity
Health status	Generalised ordinal logistic regression	Age, marital status, education, labour force status, kids in household, adults in household, household income, inadequate physical activity, smoking status, at-risk alcohol consumption, and obese
Private health insurance	Logistic regression	Age education, labour force status, kids in household, and household income Separate models developed for single persons and couples

The baseline scenario allowed the model to let equations run as set with obesity increasing in line with the trend seen over the last 20 to 30 years. The alternative scenario that removed obesity from the population provided a comparator to gain an understanding of the maximum level of change that could potentially be achieved with respect to improved health profiles and health expenditure.

The baseline scenario was run with a simulation size of 165,000. This large sample removes much of the randomness generated due to selecting samples from the census sample file on which to run the simulation. This simulation is used to gain an understanding of the background socioeconomic context within which the health module is operating. The sociodemographic modules operate exogenously to the health module. These external modules for the purpose of this exercise are taken as is and are assumed to be a correct reflection of the socioeconomic profile of Australia in 2011 and a representation of the most likely profile in 2051. The socioeconomic indicators that are used as inputs in the health module are presented in Table 4. This shows that over the 40 years between 2011 and 2051, people are more likely not to be in a marriage or de facto relationship, will have considerable increases in income, and will have a doubling of persons with bachelor degree or higher education and less people with

Specification of a logistic regression model: $\ln(P(Y_i)/P(1 - Y_i)) = \alpha + \sum_{k=1}^{K} \beta_k X_i = Z$
Probability equation for a logistic regression model: $P(Y) = \exp(Z)/(1 + \exp(Z))$
See [48]

Box 1: Logistic regression model.

Model specification:
$P(Y_i > j) = g(X\beta_j) = \exp(\alpha_j + X_i\beta_j)/(1 + [\exp(\alpha_j + X_i\beta_j)]), j = 1, 2, \ldots, M - 1$
Probability equations for the five levels of health status:
$P(Y_i = \text{Excellent Health}) = g(X_i\beta_1)$
$P(Y_i = \text{Very Good Health}) = g(X_i\beta_1) - g(X_i\beta_2)$
$P(Y_i = \text{Good Health}) = g(X_i\beta_2) - g(X_i\beta_3)$
$P(Y_i = \text{Fair Health}) = g(X_i\beta_3) - g(X_i\beta_4)$
$P(Y_i = \text{Poor Health}) = g(X_i\beta_4)$
Cumulative Distribution Function of Health Status:
Cut Point(Excellent health) = $P(Y_i = \text{Excellent Health})$
Cut Point(Very good health) = $P(Y_i = \text{Excellent Health}) + P(Y_i = \text{Very Good Health})$
Cut Point(Good health) = $P(Y_i = \text{Excellent Health}) + P(Y_i = \text{Very Good Health}) + P(Y_i = \text{Good Health})$
Cut Point(Fair health) = $P(Y_i = \text{Excellent Health}) + P(Y_i = \text{Very Good Health}) + P(Y_i = \text{Good Health}) + P(Y_i = \text{Fair Health})$
See [47]

Box 2: Generalised ordinal logistic regression model.

below year 12 education (not completing high school), and mean age will increase as does the percentage of persons aged 65 years and over.

It is important to emphasise that the results from the scenarios presented are projections not forecasts. These outcomes are unlikely to occur, particularly as health expenditure by the Australian government is highly influenced by government policy. But projections do give an indication of what could happen if certain conditions were to prevail. The projections obtained for the scenarios are highly dependent on the underlying assumptions within APPSIM related to the socioeconomic variables, whose assumed levels are discussed above.

2.5. Validation. Validation has been an ongoing process that continues as does the development and refinement of the model. For the APPSIM model, validation has primarily focussed on comparative analysis considering the simulation model outputs against external benchmark data. This has taken both a cross-sectional perspective (percentage of population in a given state) and a longitudinal perspective (looking at the rate of change or frequency of transitions) [50]. Some of the specific modules have had their validation described in detail (e.g., [51]).

Comparative validation methods were undertaken for the health module along with other aspects of validation [49]. A short description of some of the key comparative validations follows. Firstly, cross-sectional comparisons have been done. The simulation runs from 2002 out to 2051. There is actual data for the years 2001 to 2008 for some variables, with which the models outputs can be compared. In the case of obesity, there was only HILDA data for the years

2006 to 2008. In addition, the ABS National Health Surveys provided estimates of health behaviours and health status for the years 2001, 2005, and 2008 which can be used for comparison. Longitudinal aspects of the model have also been considered, both from the perspective of the trend of the data over time and whether this seems reasonable compared to alternative projections. Further, persistence of health behaviours and health status between years was calculated and compared between the actual data and the simulation model. Each of the comparisons is based on absolute differences and percentage difference, looking to produce simulation results that are within 5 percent of the benchmark data.

The health module performs well for all four health behaviours with the simulation proportions closely mirroring the outcomes estimated in HILDA. There is slightly more discrepancy between the simulation and the National Health Survey outcomes. This is a reflection of the differences between the estimates from the two surveys. There is no more than a 5 percentage point difference between the simulation outcome and the HILDA outcome for health behaviours. The average percentage point difference across the health behaviour comparisons was 1.2, a seemingly reasonable outcome. The comparative outcomes for the health status variable were not as good, with the differences ranging from less than 1 percentage point to 8 percentage points. The outcomes between the simulation and HILDA became more closely aligned as time progressed, such that from 2005, the simulation and the survey estimate were within one percentage point.

The simulation of health status performs in a manner similar to that seen for health risk behaviours. The simulation after the first couple of years starts to closely match the

TABLE 3: Models for transition probabilities within health module.

Outcome	Years	Statistical model	Explanatory variables
Inadequate physical activity	2001–2007	Pooled dynamic logistic regression	Age, couple, bachelor degree, employed, kids in household, adults in household, household income, smoker in previous year, at-risk alcohol consumption in previous year, inadequate physical activity in the previous year, and health status in previous year
Smoker	2002–2007	Pooled dynamic logistic regression	Age, couple, bachelor degree, employed, kids in household, adults in household, household income, at-risk alcohol consumption in previous year, inadequate physical activity in the previous year, and health status in previous year
At-risk alcohol consumption	2002–2007	Pooled dynamic logistic regression	Age, couple, bachelor degree, employed, kids in household, adults in household, household income, at-risk alcohol consumption in previous year, inadequate physical activity in the previous year, and health status in previous year
Obese	2006–2008	Pooled dynamic logistic regression	Age, education, current smoker, physical activity, at-risk alcohol consumption, health status, and obesity $t-1$
Health status	2006–2008	Pooled dynamic generalised ordinal logistic regression	Age, couple, bachelor degree, employed, kids in household, adults in household, household income, smoking status, at-risk alcohol consumption, physical activity, obese, and health status, $t-1$
Private health insurance	2005–2007	Multinomial logistic regression	Age, gender, and change in status of (education level, labour force status, income, household structure, and marital status)
Medical services used	2001–2008	Logistic regression model and probability tables	Age, sex, education, labour force status, health status, and private health insurance
Prescriptions used	2001–2008	Logistic regression and negative binomial regression models	Age, sex, education, labour force status, income, health status, and age-sex interactions
Hospital admissions		Logistic regression and negative binomial regression models	Age, sex, labour force status, and private health insurance
Medical costs		Tables of costs per service	Age and sex
Prescription costs		Tables of costs per script	Age and sex
Hospital costs		Tables of cost per bed day	Age

HILDA data. From 2005 there is less than one percentage-point difference between the simulation and the HILDA data. However, there is greater discrepancy present between the simulation and the Australian National Health Survey (see Figure 1).

The other key outcome considered in this paper is government health expenditure. Validation graphs for the three key sectors—medical, pharmaceutical, and hospital—that contribute to health expenditure are presented in Figure 2. The simulation outcomes are compared to administrative data supplied by Australian Department of Health and Ageing and the Australian Institute of Health and Welfare. The simulation performs adequately with most of the comparison points falling within 15 percent of the administrative data.

Beyond the cross-sectional comparative validation, some longitudinal comparisons were made for the health behaviours and health status. In Table 5, the propensity for an individual to remain in the same state between years was considered. The comparison between HILDA and the simulation show very close alignment with the worst performance being a 5 percentage point difference for inadequate physical activity. Health status was explored further looking at the transitions between the five categories. The year-on-year transitions across all health behaviours and health status are quite good in their replication of the transitions seen in HILDA data (see Table 6). In most cases, there is less than 1 percentage point difference in the estimated and simulated transitions between categories.

From the broader health perspective, mortality, which is modelled within another module of APPSIM, is presented as well. Simulated mortality rates track very closely with the historical data and ABS projections (series B) [53] as illustrated in Figure 3. However, mortality within APPSIM is currently not linked to the health module but functions separately, with deaths based on age-sex characteristics only. The simulation of mortality implicitly includes issues such as changing health status and increased life expectancy but

TABLE 4: APPSIM assumptions for input variables.

| Input variable		2011	2051
Marital status	Registered marriage	47.5	42.5
	De facto	7.7	
	Single		9.8
		44.8	
			47.7
Labour force status	Employed FT	40.2	37.9
	Employed PT	20.5	19.4
	Unemployed	3.9	3.7
	NILF	35.3	38.9
Household income (mean)		$84,700	$771,600
Number of children in household (mean)		1.2	1.1
Number of adults in household (mean)		2.2	2.2
Base scenario education levels	Bachelor degree of higher	21.5	41.1
	Diploma	7.6	8.1
	Certificate	18.0	16.7
	Year 12	14.9	15.3
	Less than year 12	38.0	18.8
Age (mean)		37.5	42.5
Proportion aged 65 years and over		13.6	22.6
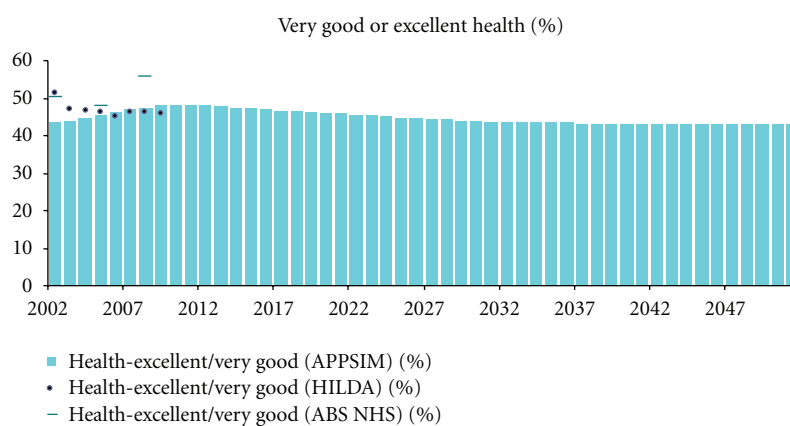
FIGURE 1: Health status validation.

TABLE 5: Percentage of population with status unchanged.

Outcome	HILDA	Health module	Difference
Inadequate physical activity	76	81	5
At-risk alcohol consumption	86	89	3
Current smoker status	95	96	1
Obesity status	92	93	1
Health status	60	60	0

currently does not explicitly model the heterogeneity of death rates associated with diseases, health behaviours, and the like. This is a major limitation of the current modelling.

3. Simulation Outcomes: Does the Obesity Epidemic Make a Difference?

The simulation results presented provided an illustrative counterfactual. The health module has been built to be expansive in nature, and the scenarios presented are to primarily provide insight into the operation of the health component of the APPSIM model. The results illustrate the mechanisms operating within the health module of APPSIM between obesity, health status, and health expenditure. The simulated trends will be considered first followed by the distributional analysis comparing the years 2011 and 2051.

TABLE 6: Year-on-year percent change for health status levels, 2006–2008.

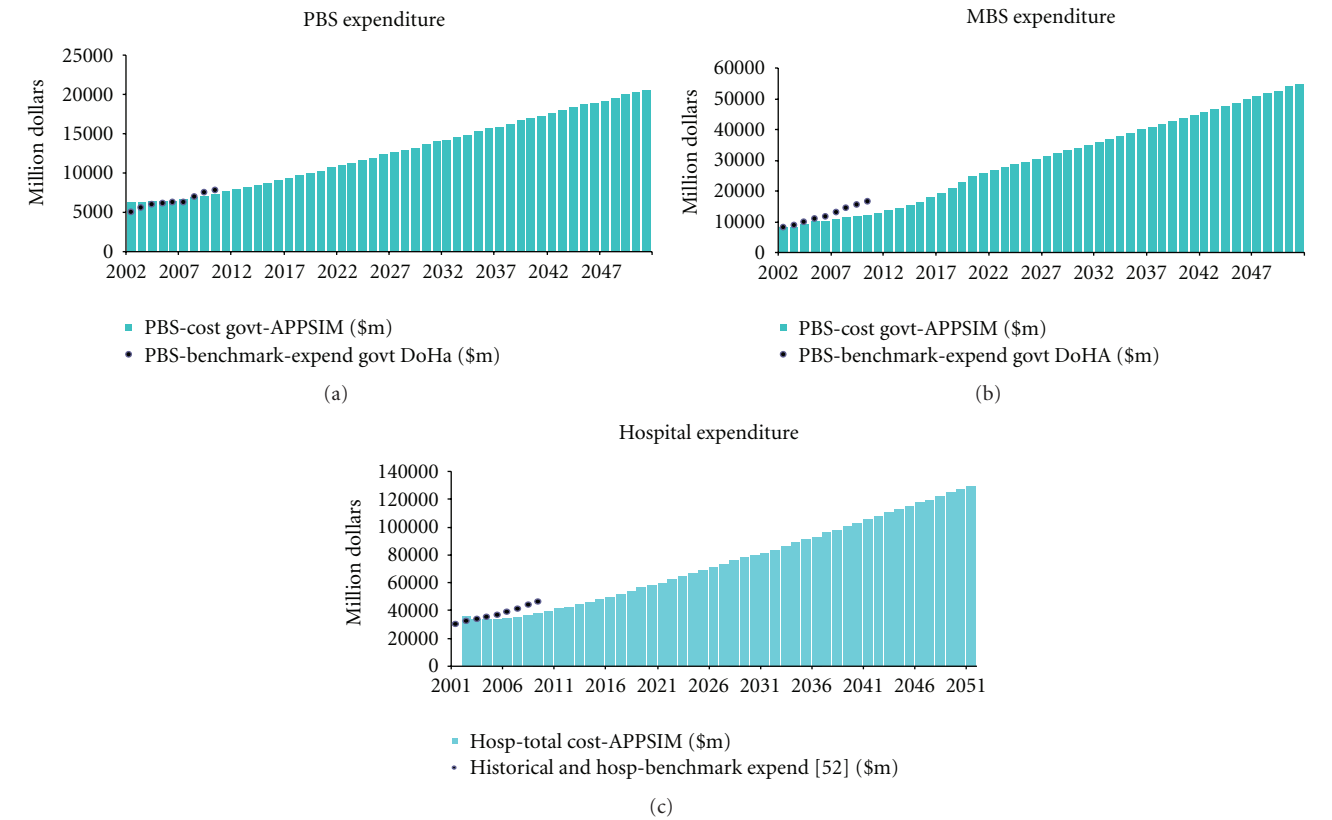| HILDA (APPSIM)	Health status at $t-1$					Total
Health status at time t	Excellent	Very good	Good	Fair	Poor	
Excellent	6.0 (6.9)	4.2 (4.1)	0.5 (0.7)	0.0 (0.1)	0.0 (0.0)	10.7 (11.8)
Very good	4.2 (4.1)	22.4 (19.8)	8.0 (7.5)	0.6 (0.8)	0.0 (0.1)	35.3 (32.4)
Good	0.6 (0.6)	8.6 (8.3)	22.7 (21.7)	4.4 (4.8)	0.2 (0.2)	36.5 (35.6)
Fair	0.1 (0.1)	0.7 (0.4)	4.8 (5.3)	7.6 (8.5)	1.1 (1.5)	14.3 (15.6)
Poor	0.0 (0.1)	0.1 (0.0)	0.2 (0.2)	1.1 (1.5)	1.8 (2.8)	3.2 (4.6)
Total	10.9 (11.8)	36.0 (32.6)	36.2 (35.4)	13.8 (15.7)	3.1 (4.5)	100

(a)

(b)

(c)

FIGURE 2: Health expenditure validation.
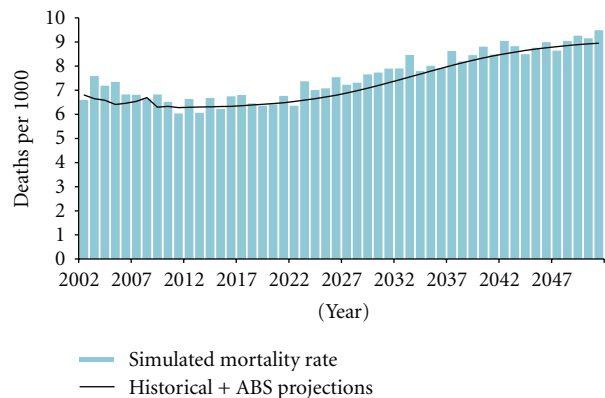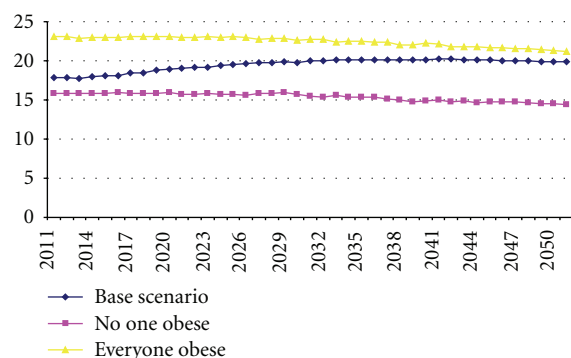

FIGURE 3: Mortality rates validation.

Figure 4: Projected percentage of the Australian population in fair or poor health, 2011–2051.

Figure 5: Projected government health expenditure in Australia, 2011–2051.

3.1. Time Series: The Big Picture. Within APPSIM, the simulated percentage of persons in fair and poor health is approximately 17 percent in 2011. This indicator drifts up to approximately 20 percent by 2030, remaining at this level for the rest of the simulation (see Figure 4). Compared to the baseline scenario, the health indicator for the simulation where no one is obese results in a lower simulated percentage of people being in fair or poor. Over the time period of 2011 and 2051, there is a slight decrease in the proportion of people in fair or poor health under the scenario of no obesity in the population compared to increasing percentage of poor health in the baseline scenario. The simulation output commences in 2002, and there are almost 10 years of high levels of obesity by 2011. The "all of the population obese" in 2011 has already a greater simulated percentage of people in fair or poor health compared to the other scenarios, and this is maintained over time. By 2051, the simulated percentage of people in fair or poor health is very similar between the baseline scenario and the "everyone obese" scenario. There is slight decrease in the "everyone obese" scenario in the later years of the projection due to the underlying impacts of other health behaviours and socioeconomic characteristics. A similar level of decline is also seen in the no-obesity scenario. Across the 2011 to 2051 time period, the simulated percentage in fair or poor health under the no-obesity scenario remains less than the other two scenarios.

The impacts on health expenditure by the government in all scenarios are large increases in health expenditure (see Figure 5). This is a reflection of the health inflation factor that is part of the initial modelling. The base scenario and the "everyone obese" scenario track in a similar manner, at a similar level, over time, with the "everyone obese" scenario always being just above the baseline expenditure. The scenario with no obesity in the population consistently tracks at lower levels than the other two scenarios. In 2051, the "no-obesity" scenario produced a simulated costs saving of almost 20 million dollars ($Aus) compared to the baseline, whilst the "everyone obese" scenario estimates an additional 6 million dollar spent. Across the years 2011 to 2051, the "no-obesity" scenario estimates an accumulated saving of nearly 400 million dollars (an average of nearly 10 million dollar each year). In contrast, the "everyone obese" scenario
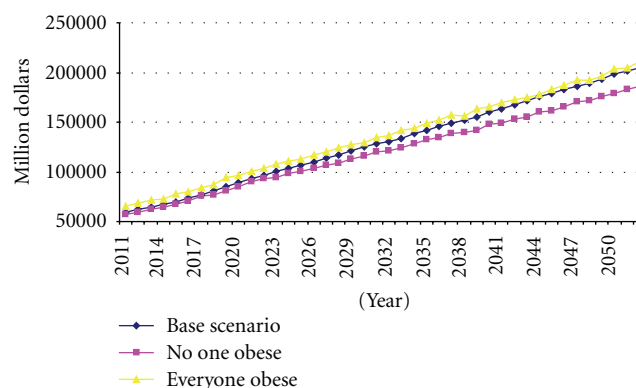
estimates an accumulated increase in spending of almost 250 million dollars (an average of approximately 6.5 million dollars each year).

3.2. Cross-Sectional Comparisons: Impacts on Subpopulations. The impacts on subpopulations across time were considered to see if any particular subgroups received additional or reduced burden of ill health. The subgroups considered were related to age groups, gender, education, and income. The age groups considered were the young (20–29 years) working age (40–49 years) and older ages (65–74 years). For education, the distribution of health status was contrasted between those with less than a year 12 education and those with at least a bachelor degree. Finally, household equivalised income quintiles were used contrasting the highest and lowest income quintiles. The modified OECD method [54] which accounts for household size and mix of adults and children was used. For this analysis, the baseline scenario and the "no one obese" scenario were considered as they offered the most contrast.

The simulation shows that between the scenarios, a greater proportion of the older cohorts were in excellent health than seen in the baseline simulation for 2051. In the baseline between 2011 and 2051, there is a general decrease in the percentage of the population in excellent health. In contrast, in the "no-obesity" scenario, there is a marginal increase in the percentage of the population in excellent health.

Across the genders, in both the baseline and "no-obesity" scenario in 2011, a slightly higher percentage of males are in excellent health compared to females. By 2051, this has reversed with slightly more females in excellent health (and slightly less females in poor health). In the "no-obesity" scenario between 2011 and 2051, both genders were more likely to be in good health compared to the baseline scenario.

The difference in health profiles is quite distinct between the education levels as seen in Figure 6. In the baseline scenario in 2011, those with least education had a lower percentage in excellent health. By 2051, this difference remains but was present to a lesser level. In the "no-obesity" scenario, by 2051 there were a greater percentage of persons in excellent health in both education groups, but the difference
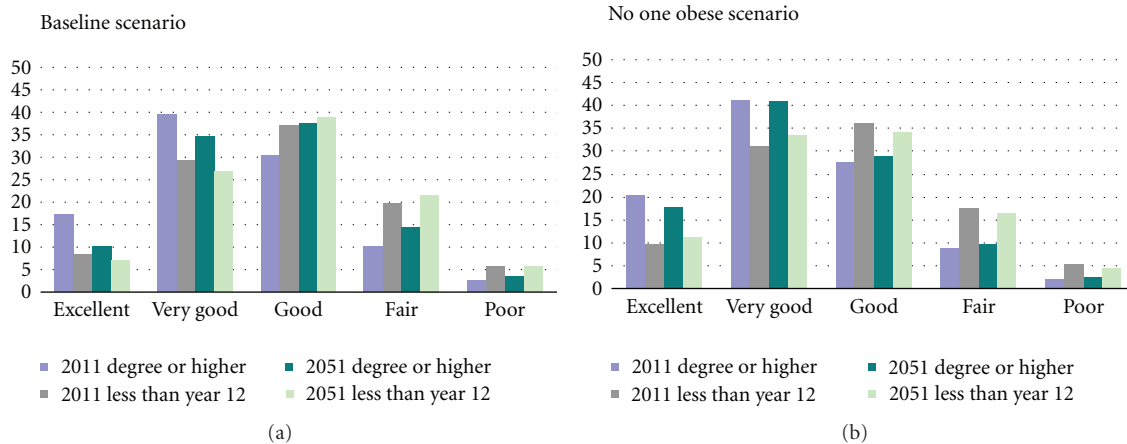
Baseline scenario

No one obese scenario

(a)

(b)

FIGURE 6: Simulated health profiles by highest educational qualification, 2011 and 2051.

between the groups remained quite large as those with lower education had less gain in health over the 40-year period.

Health differentials between income groups were similar to those seen in education; in 2011, those in the highest income quintile were in better health than those in the lowest income quintile. By 2051, the differential between the two income groups had lessened. In the "no-obesity" scenario, there was virtually no change in the proportion of persons in excellent health in either income group across the 40-year time period.

The impact on total government expenditure was considered to see if large government outlays where seen for any specific groups. In the baseline scenario, the expenditure on the oldest age cohort in 2011 is less than the expenditure on the youngest cohort in 2051. Under the "no-obesity" scenario, there is a flattening of expenditure levels across the age groups and also across time. Across both scenarios, in both time points, there is more health expenditure on females compared to males. The level of spending is less in the "no-obesity" scenario and presents less difference between the two groups across time.

The distribution of health expenditure, by educational qualification, shows a reversal of spending from 2011 where the greater levels of expenditure are on those with less than year 12 education, whilst by 2051 the greater level of expenditure is on persons with at least a bachelor degree (see Figure 7). This occurs in both scenarios but is more muted in the "no-obesity" scenario. The key reason for this reversal is the decreased numbers of person with less than year 12 education over the next 40 years.

The expenditure on those in the highest income quintile is less than that for the lowest income quintile in both scenarios. However, there is a larger differential in the level of spending between the two groups across time in the baseline scenario compared to the "no-obesity" scenario.

4. Discussion

The simulation shows that under the modelling assumptions within APPSIM, there is little additional impact on either
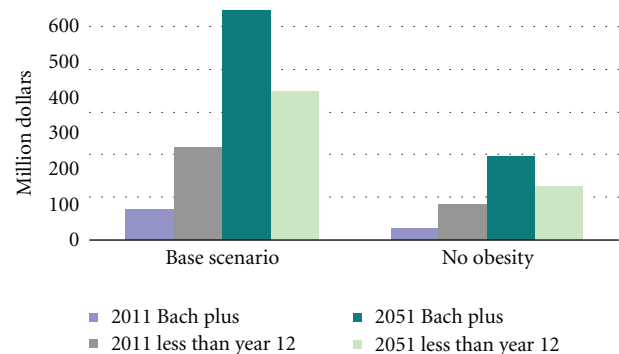

FIGURE 7: Simulated government expenditure by highest educational qualification, 2011 and 2051.

population health status or health expenditure between extremely high levels of obesity and the baseline scenario, which had a gradual increase in obesity. With the removal of obesity from the population, there is both a substantial increase in population health and some moderation of the levels of projected health expenditure. However, under each scenario, the substantial rise in health expenditure continues due to the greater influence of health cost inflation associated with expensive new technology and medications.

In all recent projections of health expenditure in Australia, regardless of the different methodologies used, there were large increases in health expenditure projected. The level of increase in health expenditure found in our simulation at $Au 200 billion spent in 2050 is more modest than that projected by the Australian Treasury at $Au 250 billion [2]. Treasury estimated this to be approximately 7 percent of GDP. Modelling by the Productivity Commission had produced estimates that health expenditure will be an even greater proportion of GDP at 10.3 percent by 2044/45 [9]. Goss in his projections, which included residential aged care within health expenditure, estimated health expenditure to have risen to $246 billion by 2033 [10]. This is a much faster rate than any of the other projections, due in part to the inclusion of the aged care costs but also the assumptions

around disease rates, proportion of disease cases that are treated, and the volume of services provided per disease case. Both the Treasury and Productivity Commission projections are sophisticated forms of cell-based modelling that primarily model health expenditure based on age and sex profiles. The modelling in our simulation has been based on age, sex, and various socioeconomic characteristics that impact on health behaviours, health status, and use of health services. However, our modelling has not pursued the simulation of diseases at the individual level over time due to lack of background health data on the basefile.

In Australia, cost impacts of obesity have been estimated (e.g., [55, 56]) showing that obesity imposed substantial direct health costs on individuals. These studies did not extend the analysis to consider the future impacts of these costs, and whether addressing the prevalence of obesity would impact the expected increases in Australian health expenditure. The impacts of obesity on health expenditure was considered by Stewart et al. [57] narrowly focussing on the cost related to cardiovascular disease. Their findings were that excess CV hospitalisations will cost $147 million per annum over the next 20 years. Our study has looked more broadly at the impacts of obesity on health status and overall health expenditure including medical, hospital, and pharmaceutical. Internationally, Wang et al. [58] looked at the impacts of obesity on health expenditure in the USA finding that for each decade out to 2030 health expenditure attributable to overweight and obesity would double. This research was extended to consider UK in 2011, showing similar results of increasing medical costs associated with obesity and related chronic diseases [59]. Our study did not consider the costs directly attributable to obesity but simulated the shift in overall health spending in Australia and the impact of shifting obesity levels on spending levels.

4.1. Limitations. As with all projections, there are assumptions that need to be made. Simulation results need to be viewed with suitable caution. To produce a functioning dynamic microsimulation model, some heroic assumptions needed to be made. The basefile did not readily lend itself to health modelling, requiring baseline imputation of health, and the transitional modelling. The need for large amounts of imputation onto the basefile meant that binary outcome variables were used. The use of such variables, whilst facilitating the modelling process, does lead to a loss of information.

Longitudinal data is limited in Australia, meaning that the development of the transition equations used a short series of panel data. The panel data used showed relatively stable health outcomes, meaning time trends needed to be applied on top of the transition model. Also, cohort effects that may be present could not be modelled due to the short data series. The modelling assumed that the relationship for certain age groups will be the same over time. That is to say, 20 year olds in 2040 are assumed to have the same behavioural responses as 20 year olds in 2006. It is highly likely that different cohorts, who have had different life experiences, will not behave in exactly the same manner. However, in the face of no alternative information, it is the "best" estimate available of what may happen in the future.

A lack of direct links between the health module and mortality means that impacts of changing health profiles (and health behaviours) on life expectancy are not modelled directly, and health expenditure impacts due to changed health profiles are diluted. In the modelling process for health behaviours and health status, it was evident that there is a dominance of the lagged variables in the transition process. That is, it takes quite large shocks to facilitate changes to individual behaviour. Consequently, it requires major changes to population characteristics, such as the complete removal of obesity from the population to start to evidence changes in the associated health variables.

Specifically considering obesity, the lack of direct links of obesity to health expenditure, and indirect links via specific obesity related diseases, meant that the possible impacts of changing obesity levels are an underestimation. However, even with this limitation, the removal of obesity from the population does result in a slowing of the escalation of health expenditure in Australia. Further development of the model will include refinement of the health outcomes considered and their linkages.

5. Conclusion

This paper has described the current structure of the health module in APPSIM, some of the validation of the module, and provided an example of policy application using the model. Health expenditure in Australia is escalating. There is potential for this to be slowed through modification of health risk factors, such as obesity. However, the impacts of ageing and expensive medical technology are significant resulting in continued healthy expenditure increases into the future.

The model will require further development of more detailed health modelling. However, the use of dynamic microsimulation for health expenditure projections has allowed the combining of data sources where one data source has not been able to provide all the information of interest with respect to the health system. Further, it has allowed distributional analysis of projections under different scenarios that has not been available in other Australian health expenditure projections.

References

[1] ABS, *Population by Age and Sex, Australian States and Territories*, ABS, Canberra, Australia, 2010.

[2] Commonwealth Department of Treasury, "Intergenerational Report 2010 Australia to 2050: future challenges," Tech. Rep., Commonwealth of Australia, Canberra, Australia, 2010.

[3] AIHW, "Health Expenditure Australia 2006-07 Health and Welfare Expenditure Series no. 35," Cat HWE 42, AIHW, Canberra, Australia, 2008.

[4] J. F. Fries, "Aging, natural death, and the compression of morbidity," *New England Journal of Medicine*, vol. 303, no. 3, pp. 130–135, 1980.

[5] M. Kramer, "The rising pandemic of mental disorders and associated chronic diseases and disabilities," *Acta Psychiatrica Scandinavica*, vol. 62, no. 285, pp. 382–397, 1980.

[6] K. G. Manton, "Changing concepts of morbidity and mortality in the elderly population," *Milbank Memorial Fund Quarterly, Health and Society*, vol. 60, no. 2, pp. 183–244, 1982.

[7] J. M. Robine and J. P. Michel, "Looking forward to a general theory on population aging," *Journals of Gerontology*, vol. 59, no. 6, pp. M590–M597, 2004.

[8] OECD, "The drivers of public expenditure on health and long-term care: an integrated approach," Economic Studies 43, OECD, Paris, France, 2006.

[9] Productivity Commission, Economic Implications of an Ageing Australia, *Productivity Commission Research Report*, Productivity Commission, Canberra, Australia, 2005.

[10] J. Goss, "Projection of Australian health care expenditure by disease, 2003 to 2033 Cat," HWE 43, AIHW, Canberra, Australia, 2008.

[11] ABS, "National Health Survey: summary of results, Australia 2004–05," ABS Cat 4364.0., ABS, Canberra, Australia, 2006.

[12] AIHW, "Australia's Health 2010 Australia's health series no. 12," Cat AUS 122, AIHW, Canberra, Australia, 2010.

[13] R. Fransoo, P. Martens, H. Prior et al., *Adult Obesity in Manitoba: Prevalence, Associations, & Outcomes*, Manitoba Centre for Health Policy, Winnipeg, Canada, 2011.

[14] K. M. Flegal, B. I. Graubard, D. F. Williamson, and M. H. Gail, "Excess deaths associated with underweight, overweight, and obesity," *Journal of the American Medical Association*, vol. 293, no. 15, pp. 1861–1867, 2005.

[15] D. L. McGee and P. Diverse, "Body mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies," *Annals of Epidemiology*, vol. 15, no. 2, pp. 87–97, 2005.

[16] K. M. Flegal, B. I. Graubard, D. F. Williamson, and M. H. Gail, "Cause-specific excess deaths associated with underweight, overweight, and obesity," *Journal of the American Medical Association*, vol. 298, no. 17, pp. 2028–2037, 2007.

[17] M. Lenz, T. Richter, and I. Muhlhauser, "The morbidity and mortality associated with overweight and obesity in adulthood: a systematic review," *Deutsches Ärzteblatt international*, vol. 106, no. 40, pp. 641–648, 2009.

[18] E. E. Calle, C. Rodriguez, K. Walker-Thurmond, and M. J. Thun, "Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. Adults," *New England Journal of Medicine*, vol. 348, no. 17, pp. 1625–1638, 2003.

[19] A. R. Dyer, J. Stamler, D. B. Garside, and P. Greenland, "Long-term consequences of body mass index for cardiovascular mortality: the Chicago heart association detection project in industry study," *Annals of Epidemiology*, vol. 14, no. 2, pp. 101–108, 2004.

[20] C. J. Lavie, R. V. Milani, and H. O. Ventura, "Obesity and cardiovascular disease. Risk factor, paradox, and impact of weight loss," *Journal of the American College of Cardiology*, vol. 53, no. 21, pp. 1925–1932, 2009.

[21] B. K. Jacobsen, I. Njølstad, I. Thune, T. Wilsgaard, M. L. Løchen, and H. Schirmer, "Increase in weight in all birth cohorts in a general population: the Tromsø Study, 1974–1994," *Archives of Internal Medicine*, vol. 161, no. 3, pp. 466–472, 2001.

[22] Juhaeri, J. Stevens, D. W. Jones, and D. Arnett, "Associations of aging and birth cohort with body mass index in a biethnic cohort," *Obesity Research*, vol. 11, no. 3, pp. 426–433, 2003.

[23] E. Dal Grande, T. Gill, A. W. Taylor, C. Chittleborough, and P. Carter, "Obesity in South Australian adults—prevalence, projections and generational assessment over 13 years," *Australian and New Zealand Journal of Public Health*, vol. 29, no. 4, pp. 343–348, 2005.

[24] S. G. Leveille, C. C. Wee, and L. I. Iezzoni, "Trends in obesity and arthritis among baby boomers and their predecessors, 1971–2002," *American Journal of Public Health*, vol. 95, no. 9, pp. 1607–1613, 2005.

[25] G. Hugo, A. W. Taylor, and E. Dal Grande, "Are baby boomers booming too much? An epidemiological description of overweight and obese baby boomers," *Obesity Research and Clinical Practice*, vol. 2, no. 3, pp. 203–214, 2008.

[26] J. Crowle and E. Turner, "Childhood obesity: an economic perspective," Staff Working Paper, Productivity Commission Staff Working Paper, Melbourne, Australia, 2010.

[27] House of Representatives Standing Committee on Health and Ageing, *Weighing It Up: Obesity in Australia*, Commonwealth of Australia, Canberra, Australia, 2009.

[28] National Preventative Health Taskforce, *Australia: the Healthiest Country by 2020*, National Preventative Health Strategy, the Roadmap for Action, Canberra, Australia, 2009.

[29] G. Orcutt, "A new type of socio-economic system," *Review of Economics and Statistics*, vol. 58, no. 2, pp. 773–797, 1957.

[30] C. O'Donoghue, "Dynamic microsimulation: a methodological survey," *Brazilian Electronic Journal of Economics*, vol. 4, no. 2, 2001.

[31] M. Spielauer, *A Dynamic Socio-demographic microsimulation model for Austria. General framework and an application for educational projections*, Ph.D. thesis, University of Vienna, Vienna, Austria, 2003.

[32] A. Scott, "Implementing health in the SAGE model," SAGE Technical Note 13, ESRC SAGE Research Group, the London School of Economics, London, UK, 2004.

[33] H. J. Smolen, D. J. Cohen, G. P. Samsa et al., "Development, validation, and application of a microsimulation model to predict stroke and mortality in medically managed asymptomatic patients with significant carotid artery stenosis," *Value in Health*, vol. 10, no. 6, pp. 489–497, 2007.

[34] C. M. Rutter and J. E. Savarino, "An evidence-based microsimulation model for colorectal cancer: validation and application," *Cancer Epidemiology Biomarkers and Prevention*, vol. 19, no. 8, pp. 1992–2002, 2010.

[35] Statistics Canada, "Health models," Ottawa, Canada, 2010, http://www.statcan.gc.ca/microsimulation/health-sante/health-sante-eng.htm.

[36] L. Flood, "SESIM: a Swedish micro-simulation model," in *Simulating an Ageing Population: A Microsimulation Approach Applied to Sweden*, P. E. S. Badi Baltagi and Wildasin, Eds., vol. 285 of *Contributions to Economic Analysis*, pp. 55–83, Emerald Group Publishing Limited, 2008.

[37] B. P. Will, C. Le Petit, J. M. Berthelot, E. M. Tomiak, S. Verma, and W. K. Evans, "Diagnostic and therapeutic approaches for nonmetastatic breast cancer in Canada, and their associated costs," *British Journal of Cancer*, vol. 79, no. 9-10, pp. 1428–1436, 1999.

[38] B. P. Will, J. M. Berthelot, K. M. Nobrega, W. Flanagan, and W. K. Evans, "Canada's population health model (POHEM): a tool for performing economic evaluations of cancer control interventions," *European Journal of Cancer*, vol. 37, no. 14, pp. 1797–1804, 2001.

[39] W. M. Flanagan, C. Le Petit, J. M. Berthelot, K. J. White, B. A. Coombs, and E. Jones-McLean, "Potential impact of population-based colorectal cancer screening in Canada," *Chronic Diseases in Canada*, vol. 24, no. 4, pp. 81–88, 2003.

[40] D. Manuel, W. Flanagan, M. Tuna, A. Okhmatovskaia, P. Finès, and C. Bennett, "Coronary heart disease risk factors in Canada: a Microsimulation predictive model," in *Proceedings of the 3rd General Conference of the Inetrnational Microsimulation Association, Microsimulation and Policy Design*, Stockholm, Sweden, June 2011.

[41] J. A. Kopec, E. C. Sayre, W. M. Flanagan et al., "Development of a population-based microsimulation model of osteoarthritis in Canada," *Osteoarthritis and Cartilage*, vol. 18, no. 3, pp. 303–311, 2010.

[42] Ministry of Health and Social Affairs, *The Future Need for Care, Results from the LEV Project*, Government Office of Sweden, Stockholm, Sweden, 2011.

[43] S. Kelly, *A Detailed Guide to APPSIM: A Description of the APPSIM Model, Its Functions and the Data Produced*, Univeristy of Canberra, Canberra, Australia, 2011.

[44] N. Watson, Ed., *HILDA User Manual—Release 9*, Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Melbourne, Australia, 2011.

[45] A. Abello and L. Brown, "MediSim (static microsimulation model of the Australian pharmaceutical benefits scheme)," in *Modelling Our Future: Population Ageing, Health and Aged Care, Model 18*, A. Gupta and A. Harding, Eds., International Symposia in Economic Theory and Econometrics, Volume 16, Elsevier B. V., Amsterdam, The Netherlands, 2007.

[46] L. Brown, A. Abello, and L. Thurecht, "Length of hospital stay by older Australians: bed-blocking or not?" NATSEM Working Paper 11/08 NATSEM Working Paper 11/08, National Centre of Social and Economic Modelling, Canberra, Australia, 2011.

[47] R. Williams, "Generalized ordered logit/partial proportional odds models for ordinal dependent variables," *Stata Journal*, vol. 6, no. 1, pp. 58–82, 2006.

[48] A. Agresti, *An Introduction to Categorical Data Analysis*, John Wiley & Sons, Hoboken, NJ, USA, 2nd edition, 2007.

[49] S. Lymer, *More than just population ageing: an evaluation using dynamic microsimulartion of the escalation of Australian Government expenditure on health over the next 40 years*, Ph.D. thesis, University of Canberra, Canberra, Australia, 2011.

[50] A. Harding, M. Keegan, and S. Kelly, "Validating a dynamic population microsimulation model: recent experience in Australia," *International Journal of Microsimulation*, vol. 3, no. 2, pp. 46–64, 2011.

[51] M. Keegan, *A dynamic microsimulation approach to simulating the impact of the labour force on issues relating to the ageing population*, Ph.D. thesis, University of Canberra, Canberra, Australia, 2010.

[52] AIHW, "Health Expenditure Australia 2009-10, Health and welfare expenditure series no. 46," Cat HWE 55, AIHW, Canberra, Australia, 2011.

[53] ABS, "Population Projections, Australia, 2006 to 2101," Cat 3222.0, ABS, Canberra, Australia, 2008.

[54] ABS, "Household income and income distribution, 2007–08," Cat 6523.0, ABS, Canberra, Australia, 2009.

[55] Access Economics, *The Growing Cost of Obesity in 2008*, Report to Diabetes Australia, 2008.

[56] S. Colagiuri, C. M. Y. Lee, R. Colagiuri et al., "The cost of overweight and obesity in Australia," *Medical Journal of Australia*, vol. 192, no. 5, pp. 260–264, 2010.

[57] S. Stewart, G. Tikellis, C. Carrington, K. Walker, and K. O'Dea, *Australia's Future "Fat Bomb": A Report on the Long-Term Consequences of Australia's Expanding Waistline on Cardiovascular Disease*, BHRI, Melbourne, Australia, 2008.

[58] Y. Wang, M. A. Beydoun, L. Liang, B. Caballero, and S. K. Kumanyika, "Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic," *Obesity*, vol. 16, no. 10, pp. 2323–2330, 2008.

[59] Y. Wang, K. McPherson, T. Marsh, S. Gortmaker, and M. Brown, "Health and economic burden of the projected obesity trends in the USA and UK," *The Lancet*, vol. 378, no. 9793, pp. 815–825, 2011.

[60] Department of Health and Ageing, *National Physical Activity Guidelines for Adults*, Commonwealth of Australia, Canberra, Australia, 1999.

[61] National Health and Medical Research Council, *Australian Guidelines to Reduce Health Risks from Drinking Alcohol*, Commonwealth of Australia, Canberra, Australia, 2009.

[62] WHO, "Fact Sheet No. 311: Obesity and overweight," 2011, http://www.who.int/mediacentre/factsheets/fs311/en/.