CENTRE FOR RESEARCH IN ECONOMIC SOCIOLOGY AND INNOVATION
WORKING PAPER SERIES: ISSN 1759-2860





An Agent-based Model of Household Spending Using a Random Assignment Scheme

CRESI WORKING PAPER NUMBER: 2012-03

Tony Lawson tlawso@essex.ac.uk

Centre for Research in Economic Sociology and Innovation
Department of Sociology
University of Essex
Colchester
Essex
UK

This paper was presented at the 3rd General Conference of the International Microsimulation Association, 8-10 June 2011, Stockholm, Sweden.

University of Essex

Abstract:

This paper describes a way to model household expenditure using an agent-based microsimulation approach. Work in this area is dominated by econometric models, which attempt to estimate the parameters of an equation linking household characteristics with expenditure patterns. However, as Klevmarken (1997) notes, the representation of behaviour in these models is limited by the current state of the art in economics. A way to overcome these limitations is proposed based on a random assignment scheme where data is obtained from a donor case that is in some way similar to the receiving unit.

The paper begins with a brief account of methods for modelling behaviour currently used in microsimulation. It then introduces the random assignment scheme in an example describing how household incomes can be projected over time. This is illustrated by describing two models that have been implemented using the approach. One of them is to forecast changes in household expenditure in response to variations in income. The other models the effect of the level of unemployment on household spending patterns. These examples provide the basis for a discussion of the strengths and weaknesses of the random assignment met.

Keywords/tags:

random assignment, microsimulation, income, employment, expenditure

Citation:

Lawson, T., (2012) *An Agent-based Model of Household Spending Using a Random Assignment Scheme*, Centre for Research in Economic Sociology and Innovation (CRESI) Working Paper 2012-03, University of Essex: Colchester

About CRESI:

Based in the UK's leading Sociology Department, the Centre for Research in Economic Sociology and Innovation (CRESI) is the first UK centre for research in economic sociology. With a clear focus on innovation, our research programmes highlight contemporary and historical processes of socio-economic transformation. You can read about our research and join our conversation.

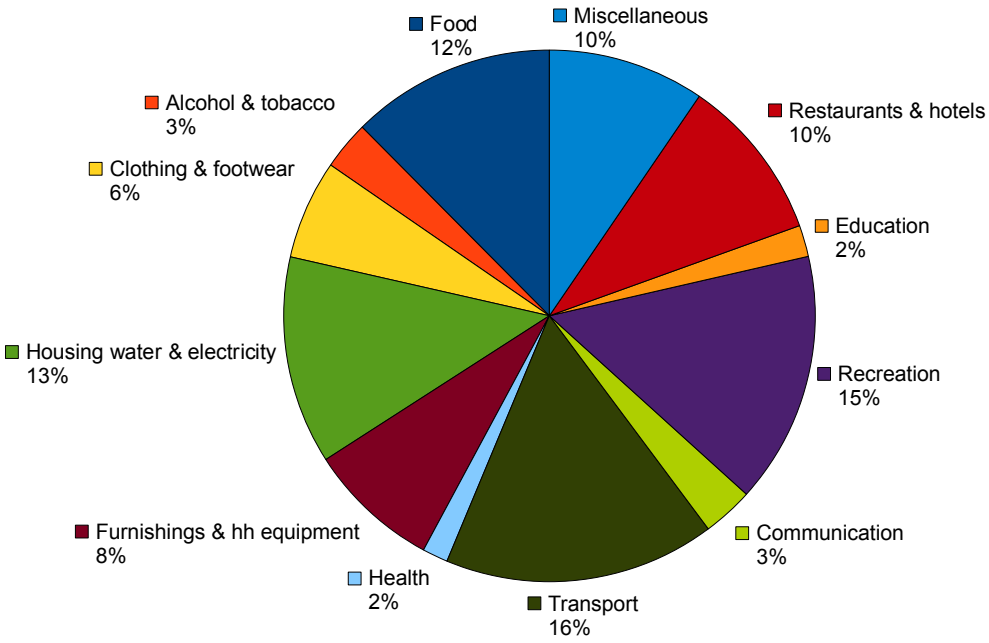
Table of Contents

1 INTRODUCTION

Consumer spending in the UK amounted to 872 billion pounds in 2009 (ONS, 2010). It is understandable therefore, that commercial and public organisations have an interest in gaining a better understanding of this sector of the economy. In the private sector, this might be to predict the size of the market for particular goods or services. For governments, it is important to understand the effect of indirect taxes that affect households differently depending on the type and quantity of goods they consume.

Figure 1 shows the proportion of household income that was spent on a range of categories as defined in the 2006 EFS (Expenditure and Food Survey).

Figure 1: Household Budget Shares



The problem that this paper addresses is to find ways to model and forecast how spending on various goods and services varies in response to demographic, economic and socio-technical change. Put another way this is, how to estimate the behavioural response of households to changes in their social and economic circumstances.

Klevmarken (1997) describes four families of behavioural modelling as they apply to microsimulation.

1. **Transition probabilities** – The behavioural response is chosen at random from a set of mutually exclusive possibilities on the basis of their likelihood or probability of occurring. This is often used in dynamic microsimulation to model demographic transitions.
2. **Count data models** – These are based on the amount of time spent in various states. This approach can be used if there are insufficient data to calculate transition probabilities.
3. **Continuous data models** – These include systems of equations or regression models as well as econometric models. They are widely used in microsimulation, particularly in the area of modelling economic behaviour.
4. **Random assignment –** Behavioural response is determined by finding a donor which is in some sense similar to the receiving unit.

According to Klevmarken, the first three types belong to the conventional econometric paradigm of estimating an average structure and then applying random deviations to it. The

fourth family are known as random assignment schemes were the structure is not estimated but is implied in variables that define 'closeness' between particular cases. The advantages of this method are that it is not necessary to impose a functional form on the data or make any assumptions about the distribution of variables. There are no parameters to estimate and the method preserves the variation and most of the correlation present in the original dataset. The approach also allows the study of situations where people behave in fundamentally different ways; in particular where some individuals do something other than maximise their utility function (Klevmarken, 1997).

Despite these advantages, random assignment schemes seem to be used rarely in microsimulation modelling. (Klevmarken et al., 1992), (Klevmarken & Olovsson, 1996) (Holm, Mäkilä & Lundevaller, 2009) are some of the few examples in the published literature. Klevmarken recommends that this approach be explored further and this paper can be seen as one such exploration as it is applied to the study of consumption. It begins with a worked example showing how a random assignment scheme can be applied to project a small panel of hypothetical income data over time. Next, the method is adapted to model household expenditure and two models are described. The first one is to predict how households respond to changes in income. The second simulates the effect of changes in the rate of unemployment on household spending patterns. Both were implemented in the form of a microsimulation using the agent-based modelling platform NetLogo. The findings from the modelling phase are then discussed in terms of the practical strengths and weaknesses of the random assignment scheme. Finally, a way to overcome one of the weaknesses of this approach is proposed which suggests a way to incorporate theoretical insights into microsimulation modelling.

2 RANDOM ASSIGNMENT SCHEME

Klevmarken (1997) describes how what he calls a random assignment scheme can be used as a method of projecting a variable over time. Data is available on a set of incomes for two consecutive years and it is desired to project the income distribution for the following year. This is done by first defining a distance metric between a donor's income in year 1 and a receiver's income in year 2. In this example, it is simply the arithmetical difference between the two. The essence of the procedure is that the income for each case in year 3 is obtained by finding a donor whose income in year 1 is similar to the current case's income in year 2. The receiver's income in year 3 is then assigned to be what the donor's income subsequently became in year 2.

The following example illustrates the process. Table 1 shows an imaginary dataset with 10 cases.

Table 1: Initial Dataset

Case Number	Income in Year 1	Income in Year 2
1	120	130
2	110	120
3	90	100
4	100	110
5	70	80
6	80	90
7	130	140
8	100	110
9	90	100

10	110	120
Mean	100	110

Case 1 has an income, in year 2, of 130 units. In year 1, case 7 had an income of 130 so it becomes the donor for case 1. Case 2 has an income of 120 in year 2 and case1 had an income of 120 in year 1 so case 1 becomes the donor for case 2. This process continues until all cases have been assigned a donor.

Table 2: Donor Case Assignment

Case Number	Income in Year 1	Income in Year 2	Donor Case Number	Donor Income in Year 1	Donor Income in Year 2
1	120	130	7	130	140
2	110	120	1	120	130
3	90	100	4	100	110
4	100	110	2	110	120
5	70	80	6	80	90
6	80	90	3	90	100
7	130	140	7	130	140
8	100	110	2	110	120
9	90	100	4	100	110
10	110	120	1	120	130
Mean	100	110	Mean	109	119

A problem is encountered when Case 7 requires a donor whose income in the first year is 140. Since it has the highest income, there is no exact match and the next lower income case is used which is case 7 itself. When all matches have been completed, the donor income of year 2 becomes the recipient's income in year 3.

Table 3: Completed Income Projection

Case Number	Income in Year 1	Income in Year 2	Income in Year 3
1	120	130	140
2	110	120	130
3	90	100	110
4	100	110	120
5	70	80	90
6	80	90	100
7	130	140	140

8	100	110	120
9	90	100	110
10	110	120	130
Mean	100	110	119

The method operates by effectively applying last year's change as a forecast of what will happen next year. This is a limitation in that, if incomes changed by a different amount each year, it would not be predictable using this method. Also, there is an anomaly in that the mean income in year 3 is 119 and not the 120 that would be expected. This arises from there being no suitable donor for case 7 and illustrates one of the limitations noted by Klevmarken (1997) which is that it is not possible to predict beyond the range of values already present in the original sample.

Despite this, it can be seen that the procedure has, for the most part, correctly implemented the implied rule that the income next year is 10 units higher than it was this year. In this example, it is easy to deduce what the rule should be, however the power of this method is that since the rule remains implicit, it is possible represent an arbitrarily complex system of rules with no additional complexity to the model. This means that a system can be modelled without first having to discover precisely how it works. In modelling aspects of human behaviour, such as consumption, where the rules are complex and may be impossible to discover, this is an important advantage.

This scheme can quite readily be turned into a method for forecasting household expenditure. Suppose there is a change in household circumstances, such as a new household member moves in. It might be expected that spending on some items like food would increase while others such as rent or mortgage would not be affected. The purpose of the microsimulation model is to determine which expenditure categories would change and by how much. This can be done by locating a household within the dataset that already has a composition that is as similar as possible to the new household and copying its expenditure pattern. The matching variables would include demographic characteristics and financial variables such as income. These variables would be chosen because they are thought to have some correlation with expenditure patterns and they play a role corresponding to the independent variables in regression modelling. It is assumed that whatever changes the household in question will make as a result of the new member arriving, will already have been made by the donor household so this behavioural response will be embedded or encoded within the expenditure pattern that is copied.

This process can be represented as follows:

```
load cross-sectional household data file
for each simulated year
        for each household
                make changes to household characteristics
        for each household
                if the household has any changes
                        locate a similar household
                        copy its expenditure pattern
calculate new aggregated expenditures for categories of interest
```


3 THE EFFECT OF CHANGES IN HOUSEHOLD INCOME ON EXPENDITURE PATTERNS

The Expenditure and Food Survey (EFS) provides data on an array of expenditure categories and this makes it suitable for use as the base data set for the model. This is an annual cross-sectional survey that collects detailed information on household spending. Its sample size is around 6000 households containing over 10,000 individuals. The EFS documentation provides a description of each expenditure category along with a list of its components. Table 4 provides a brief summary of each variable and some notes on what is included.

Table 4: Primary EFS Expenditure Categories

Variable Name	EFS Household Expenditure Category	Notes
P601t	Food & non-alcoholic Drinks	
P602t	Alcohol Tobacco & Narcotics	Alcohol to be consumed at home
P603t	Clothing & Footwear	
P604t	Housing Fuel & Power	Includes rent, maintenance of household, water and fuel bills
P605t	Household Furnishings & Equipment	Includes carpets, curtains, household appliances, utensils and tools
P606t	Health	Prescriptions, glasses, dentist fees but not medical insurance
P607t	Transport	Purchase of vehicles, fuel, vehicle maintenance but not insurance
P608t	Communications	Mobile and fixed line telephone, postage but not internet subscription
P609t	Recreation & Culture	Television, computers, CDs, boats, caravans, pets, sports, holidays
P610t	Education	Course fees, school trips
P611t	Restaurants & Hotels	Includes takeaways, alcohol consumed outside the home and school meals
P612t	Miscellaneous	Includes insurance, jewellery, child care, fees, moving expenses

The above categories together make up what the EFS calls Total Consumption Expenditure P600T. This excludes items like mortgage interest payments, council tax, motor vehicle tax, money spent abroad, gifts, interest on credit cards, income tax, central heating installation, purchase of dwellings, capital repayment of mortgage, DIY improvements and savings.

The base population was advanced over time using an existing dynamic microsimulation model (Lawson, 2009). The demographic modules were removed so that the effects of changing income could be isolated. The 12 primary expenditure categories were projected for 20 years

under the scenario of a 5% annual increase in real household income. An additional category 'Other' was defined as the difference between 'total expenditure' FSALL and the sum of the 12 consumption items. The category 'other' will then include some important items such as mortgage interest payments.

As income rises, some of it is not spent but goes on other items like savings and income tax. Since these are not considered to be consumption items, the share of total income spent on the consumption items decreases. Expressing spending as a share of expenditure avoids this effect.

The results are shown over three graphs for clarity.

Figure 2: Share of Expenditure on Food, Alcohol, Clothing and Housing

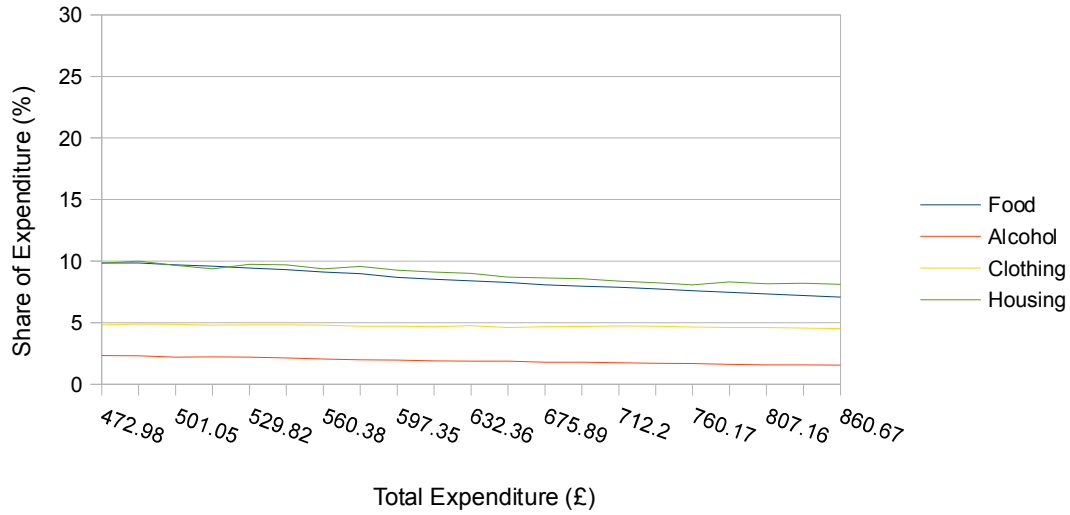


Figure 3: Share of Expenditure on Furniture, Health, Transport and Communication

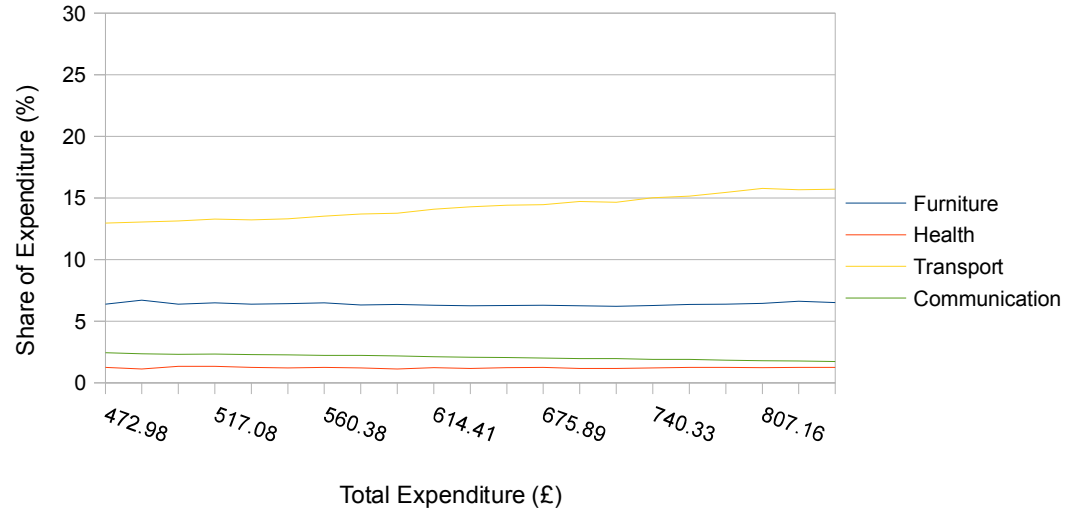


Figure 4: Share of Expenditure on Recreation, Education, Hotels, Miscellaneous and Other

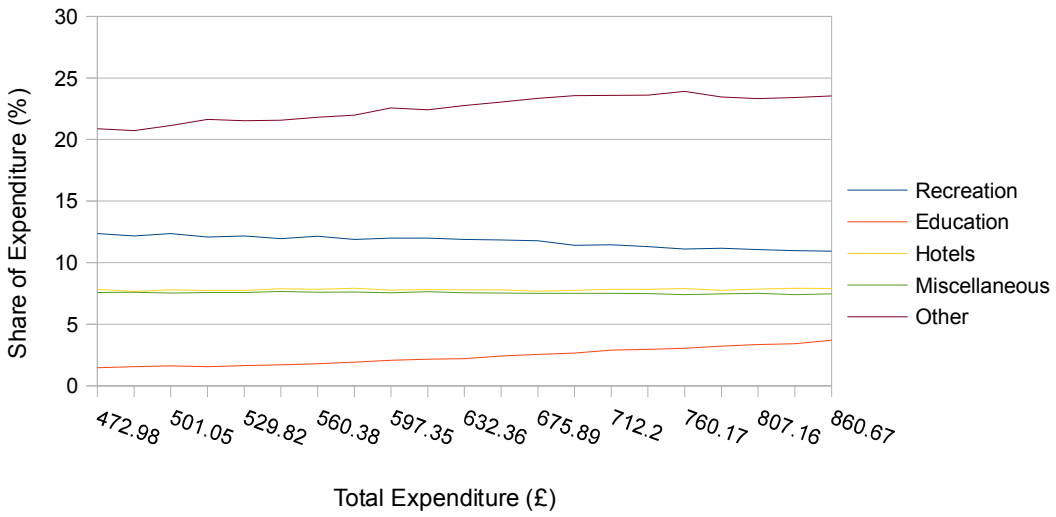


Table 5 shows how the balance of spending changes as total expenditure increases. The items at the top of the list receive a disproportionate amount of the increased expenditure. At the bottom end, spending still increases in nominal terms but decreases as a proportion of expenditure.

Table 5: Share of Expenditure with increasing Income

Expenditure Category	Initial Share (%)	Final Share (%)	Percentage Change
Education	1.47	3.7	151.25
Transport	12.97	15.72	21.23
Other	20.86	23.55	12.88
Furniture	6.38	6.5	1.92
Hotels	7.84	7.9	0.64
Health	1.25	1.23	-0.96
Miscellaneous	7.57	7.46	-1.52
Clothing	4.82	4.53	-6.07
Recreation	12.37	10.94	-11.53
Housing	9.83	8.11	-17.53
Food	9.85	7.07	-28.19
Communication	2.45	1.72	-29.48
Alcohol	2.34	1.56	-33.12

4 EMPLOYMENT MODEL

Another factor that might be thought to have an impact on household spending decisions is the availability of employment or risk of unemployment. This presents a slightly more complex challenge from a modelling perspective because unemployment operates at the individual level while household expenditure depends on the whole household. As levels of unemployment change, there needs to be some way to cascade changes in the prevalence of unemployment for individuals, to household level properties such as the number of workers or the number of unemployed within the household and from there to household spending patterns.

The rate of unemployment is varied exogenously and this causes a number of individuals selected from the population by age and sex groups to become unemployed. This has an effect on the composition of the household in which they live. Expenditure patterns are then modified by copying the income and expenditure pattern from an existing household that already has the same or similar composition.

4.1 Initial Data

The British Household Panel Survey (BHPS) (ISER, 2008) is a large scale longitudinal survey of the UK population, based on detailed interviews with over 5000 individuals in 10,000 households. While the BHPS records expenditure data for only a few utility items, it contains more detailed employment and demographic data than the EFS and this makes it suitable for use as the base dataset for the employment model.

At the individual level, the following variables are taken from the record PINDRESP in the 2006 BHPS:

Table 6: Individual Level BHPS Variables

Variable Name	Description	Model Values
PAGE	Age at date of interview	
PSEX	Sex	
PJBSTAT	Current labour force status	0 Unknown
		1 Self Employed
		2 Employed
		3 Unemployed
		4 Retired
		5 Family Care
		6 Full Time Student
		7 Long Term Sick/Disabled
		8 Other

The household level variables used in the employment model are listed in table 7.

Table 7: Household Level Variables

Variable Name	Description	Model Values

PXPMG	Last total monthly mortgage payment	
PRENTG	Gross rent including housing benefit	
PXPFOOD	Total weekly food and grocery bill	
PFIHHYR	Annual household income	
PXPLECLW	Expenditure on electricity last week	
PXPGASLW	Expenditure on gas last week	
PXPHSN	Net monthly housing costs	
PXPOIL	Monthly expenditure on oil	

Once read in, households are weighted to correspond to the UK population. This is done by duplicating or deleting households according to their household weight PXHWGHT. Then, for each year that the simulation is to run, a target rate of unemployment can be set by using a slider on the user interface or by setting a value in the program code. This is defined as the proportion of the population who are active in the labour market who are not currently either employed or self employed. The categories of individual who are considered to be inactive are 'retired', 'family care', 'full time student' and 'long term sick/disabled'. Any agents whose employment status is 'other' or 'unknown' are also considered to be inactive.

At the start of each year, the exogenous target unemployment rate is compared with the rate of unemployment in the current agent population. If there is a difference, the number of agents is calculated that will be required to change employment status to bring the current population into line with the target population. This is done using the formula:

n = numelligible / 100 * (unemployment_rate – u%)

n is the number of transitions required

if n is positive, the number of unemployed must be increased

if n is negative, some of agents currently unemployed are re-employed.

numelligible is number of agents that make up the workforce. This is the sum of those with an employment status of 'self employed', 'employed' and 'unemployed'.

unemployment_rate is the target value that is set in the program.

u% is the current percentage of eligible agents that are unemployed.

The next issue is to select which particular agents in the current working population are to become unemployed. There are many ways this could be done and for the purposes of this example, the implementation takes the form of a random process modified by age and sex groups.

According to the BHPS, the make up of the unemployed population in 1991 is as shown in Table 8.

Table 8: Make up of UK Unemployed Population

Male		**Female**	
Age Band	**Percent of Total Unemployed**	**Age Band**	**Percent of Total Unemployed**
16-29	27.1	**16-29**	15.0

30-49	24.6	**30-59**	14
50-65	15.8		
65+	2.5	**60+**	1.0
Total	**70.0**		**30**

It might be expected that the next person to become unemployed would have characteristics that fit into this distribution. There would then be a 27.1% chance of that person being a male aged 16 to 29. A 15% chance of them being a female aged 16 to 29 and a 14.0% chance of them being a female aged 30 to 59 etc. However, if we keep drawing from the employed population using these probabilities, eventually some of the groups would contain no members. For example, the unemployed population is about 70% male and 30% female. If we select using this ratio then the number of employed males would eventually reach zero and the working population would be entirely female. To ameliorate this effect, the selection takes place in two stages. First, an agent is chosen from the working population purely at random. This agent is considered as a candidate for unemployment. Next, a draw takes place according to the probabilities above. If selected, the agent is classified as unemployed. If not selected, another agent is chosen at random and the process begins again. This continues until n successful transitions have been completed.

The BHPS includes some respondents who considered themselves to be unemployed although they are aged above what is normally considered to be the age of retirement in the UK. In the model, employment transitions only apply to men aged 16 to 65 and women aged 16 to 60. This means that those above the retirement age who are categorised as employed or unemployed in the BHPS remain in that state. However, the probability of transition that would have applied to this group (2.5% for males aged 65 and over) is assigned to the next oldest group (in this case to males aged 50 to 65).

A similar process operates to move unemployed agents into employment but with probabilities as shown in Table 9.

Table 9: Make up of the UK Employed Population

Male		**Female**	
Age Band	**Percent of Total Employed**	**Age Band**	**Percent of Total Employed**
16-19	2.3	**16-19**	2.4
20-49	40.2	**30-49**	33.8
50-67	12.3	**50-67**	8.8
68+	0.2	**68+**	0.0
Total	55		45

Agents are selected at random and assigned into employment according to the above ratios until sufficient transitions have been implemented to reach the target unemployment rate. When an individual is made unemployed, a record is stored of whether they were self employed or worked at their employer's premises prior to the transition. If this individual moves back into employment, they return to their previous location of work.

Once all the individual level employment transitions have been completed, a new expenditure pattern can be assigned to the households that have experienced a change in composition resulting from a transition in the employment status of one or more household members. Any

household where the number of employed or unemployed occupants has changed during the current year will have its income and expenditure set replaced with one from a household which has the same demographic type, number of occupants, number of adults, number of children, number inactive, number employed, number unemployed and number retired. The motivation for applying such a long list of constraints is to find a household that is as similar as possible to the current household. In some of the early prototypes, only the number of unemployed was considered. However, it was found that an increase in the number of unemployed within a household could then be correlated with a larger household. This, in turn, might be correlated with a higher income that would mask much of the effect of increasing unemployment levels. The use of such stringent conditions is not without its own problems however. In some cases, there is no exact match within the population. In some prototypes this was found to be up to 18% of attempted copies. This prompted a refinement where there is a relaxation in the conditions for those households for which an exact match cannot be found. In these cases, the income and expenditure pattern is copied from a household that has the same number of occupants and the same number of unemployed. This was found to reduce the level of unmatched households to between 1 and 2 percent. Nevertheless, the simulation will underestimate the change in household expenditures due to a change in the rate of unemployment by a corresponding amount.

A scenario was run where the rate of unemployment was raised from its initial 2006 level of 4.6% to 10%, 15% and 20% after 1, 2 and 3 years respectively. Increasing the rate of unemployment leads to a decline in mean household incomes and this is represented in Table 10.

Table 10: Mean Annual Household Income under Rising Unemployment

Years	Annual Household Income (£)
0	29,950
1	28,430
2	27,630
3	27,120

The changes in mean weekly household expenditure are shown in Table 11 for the categories of food, housing, mortgage, rent and utilities. Housing is calculated as the sum of rent and mortgage. Utilities is the sum of electricity, gas and oil. In some of the categories, a significant number of households have no expenditure. This is particularly evident in the categories of housing, rent and mortgage. The zero expenditures have the effect of reducing mean expenditures to below what might be thought of as typical values.

The results are shown in Table 11 and represented graphically in Figure 5.

Table 11: Changes in Weekly Spending Under Rising Unemployment (£ per week)

Years	Food (£)	Housing (£)	Mortgage (£)	Rent (£)	Utilities (£)
0	65.95	60.79	47.19	10.46	14.94
1	64.71	56.59	42.96	12.56	14.92
2	64.09	54.27	40.50	13.55	14.88
3	63.42	52.57	38.86	14.71	14.81

From table 11 it can be seen that spending on food decreases from £65.95 to £63.42. This

represents the total value of the decrease in food consumption across the whole population due to the slightly over 15% of individuals who transitioned to unemployment. The effect on the households that were directly affected by the change would be proportionately greater.

Figure 5: Changes in Weekly Spending under Rising Unemployment

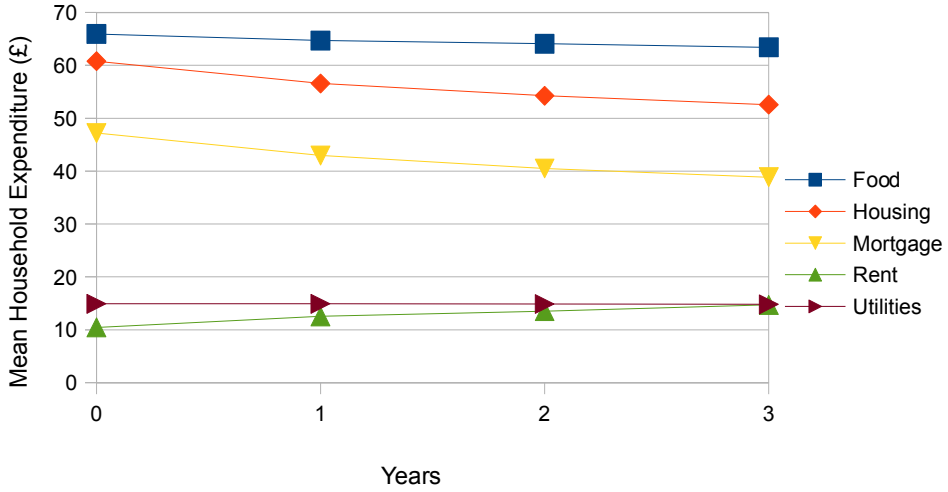


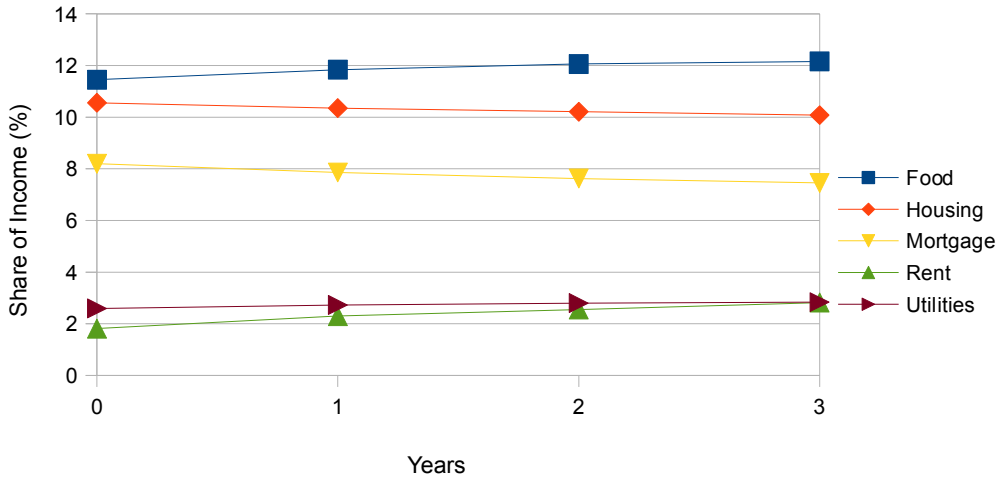
Figure 5 shows that spending on food, and mortgage falls as unemployment rises. Surprisingly, spending on rent appears to rise. A possible explanation for this might be that employment status is correlated with housing tenure in that rented accommodation contains a higher proportion of unemployed occupants. This could be controlled for by including housing tenure in the matching criteria. Spending on utilities manifests a small fall of 13p per week. Although the lower incomes brought about by increases in unemployment would presumably lead to a desire to restrict spending on utilities, the increased occupancy of the household due to unemployment may lead to some upward pressure on utility bills.

Table 12 shows the results in terms of percentage share of weekly household income.

Table 12: Changes in Share of Income under Rising Unemployment

Years	Food (%)	Housing (%)	Mortgage (%)	Rent (%)	Utilities (%)
0	11.45	10.56	8.19	1.82	2.59
1	11.84	10.35	7.86	2.30	2.73
2	12.06	10.21	7.62	2.55	2.80
3	12.16	10.08	7.45	2.82	2.84

Figure 6: Changes in Share of Income under Rising Unemployment



Although actual spending on food falls, the proportion of income spent rises as would be expected from Engel's Law (Engel, 1857, 1895) . The share of income spent on rent and utilities also rises indicating the essential nature of these items.

5 EVALUATION

While the results of these models have some interest in their own right, the main purpose of constructing then was to gain some understanding of how the random assignment scheme works in practice. Klevmarken described some of the more theoretical advantages and disadvantages associated with this method. To this can now be added some more practical insights gained during the implementation of the models.

5.1 Strengths

5.1.1 Cross-sectional

The models described in this paper take as their base data set a cross-sectional survey from one particular year. The heterogeneity of the households is used to forecast how they might change in the future. If the model relied on longitudinal data to make future predictions this would limit the applicability of the method because cross-sectional surveys are much more common than longitudinal data sets.

5.1.2 Compatible with Microsimulation

The random assignment scheme is derived from statistical matching, which is already well known within the microsimulation community. It preserves the essential characteristic of microsimulation which is the representation of individual units. Groups or classes of units do not have to be defined until they are output so that any change to the choice of groups does not affect the operation of the model itself.

5.1.3 Multi-level

The model was shown to be able to deal simultaneously with both the individual and household level. The demographic changes that drive household composition were assumed to arise from the individual level. This then affects household expenditure by a process of upward causation.

5.1.4 Flexibility

If some change is implemented in the model, increasing incomes for example, the matching procedure will copy the expenditure pattern from a household that already has a higher income. When this is done, all the variables for that case will be available for output. This

means that a model which is designed to predict the effect of income on the budget share for a particular set of items can be changed to study a different set by including them in the copy routine and in the output. The result is a powerful and flexible method which is relatively easy to change for different applications.

5.1.5 Minimal Assumptions

The essence of this method is that when a household experiences a change in circumstances, its response will be to alter its spending pattern to become more like that of a household that has been exposed to the new circumstances for some time. This is based on two assumptions firstly that similar households, on average, manifest similar spending patterns and secondly that households will adapt to changes in economic or demographic conditions by changing their spending pattern in such a way that the first assumption remains valid. These assumptions are relatively transparent and can be tested empirically provided there is a suitable dataset. They can be varied to determine their effect; for example by changing the copying mechanism so that only part of the difference in expenditures is applied. Also, they are not derived from economic theory so they provide an independent method of making projections which can be compared against those obtained by traditional econometric methods.

5.2 Weaknesses

5.2.1 Side Effects

The ability to copy all aspects of a case can create problems if it is done without regard to causation. In the employment model, a rise in the level of unemployment led to a fall in average incomes. This makes sense because in most cases, unemployment benefit is less than what was previously obtained from wages. In this sense, unemployment causes lower income. However, it would equally be possible to use the income change model to predict changes in employment. If this was done it would be seen that as incomes fall, unemployment rises. This would be a nonsense result because income, of itself, does not cause unemployment. The two are correlated but in this case, causation is being applied in the wrong direction.

5.2.2 Edge-effects

When modelling a continuous variable such as income, special consideration must be given to the cases at the upper and lower extremes. When incomes rise, there will be no case to copy from for the household that already has the highest income. When incomes fall, there will be no appropriate donor case for the household that has the lowest income. In the income change model, this was addressed by keeping budget shares constant and inflating or deflating the amount spent on each commodity in proportion to the change in income.

In practice there is no evidence that people will behave in this way and it is probable that this assumption will lead to an underestimate of the response to changes in circumstances. According to this assumption, if incomes are halved, the amount spent on food by the poorest household will also half. If incomes half again, the amount spent on food will half again. This is implausible because there is a physical minimum amount of food required for life and it might be expected that the household spending on this could not fall below a certain level. In this situation, the budget share for food would increase by more than the model would predict.

This would not be a significant problem if it only applied to the household at the end of the distribution. Unfortunately however, the household with the second lowest income may at some point copy the expenditure pattern from the lowest. The third lowest income household will copy from the second lowest and so on. In this way, the distortion introduced at the lowest income household will spread to others as they pass the point where the lowest income household was originally.

A similar problem exists in predicting how the highest income household will respond as their income increases. Here, doubling the income would result in twice as much being spent on each commodity and the model will not be able to predict spending on any new categories that were not previously budgeted for.

6 FURTHER WORK

These problems arise from the disadvantage of the random assignment scheme that Klevmarken pointed out, which is that behaviour is restricted to what has already been observed. However, it would be possible to add behaviour to these cases provided it could be deduced from another source. This could be obtained empirically or by drawing on the wealth of theoretical insights into human behaviour in fields such as psychology and sociology. It would also be possible to apply behavioural assumptions for all the agents, not just the ones at the ends of the distribution. In this way, the microsimulation would take on some of the traits of agent-based modelling. One way this can be done is to modify the household immediately before a similar household is located. The modified pseudo-code would then appear as follows:

```
load cross sectional household data file
for each simulated year
        for each household
                make changes to household characteristics
        for each household
                if the household has any changes
                        modify unit according to behavioural assumptions
                        locate a similar household
                        copy its expenditure pattern
    calculate new aggregated expenditures for categories of interest
```

In doing so, it proposes a way to merge a theory driven agent-based model with a probabilistic microsimulation model.

7 CONCLUSION

Random assignment schemes have been used rarely compared to more established econometric systems. The research described in this paper has demonstrated the feasibility of a random assignment scheme as a way of modelling the effect of household income and employment on consumption patterns. It found that in practice, random assignment is a flexible method of modelling household consumption behaviour that preserves the essential characteristic of microsimulation which is to work with individual cases. Random assignment schemes do not rely on the assumptions of economic theory and as a result this gives scope to model the full heterogeneity of human behaviour. Further work is indicated to improve the validation of this approach and integrate microsimulation modelling with more theoretically driven approaches.

8 ACKNOWLEDGEMENTS

The work reported here was supported by the UK Economic and Social Research Council (ESRC) and BT via a Collaborative Research Studentship "Methods and Tools for the Microsimulation and Forecasting Household Expenditure".

The British Household Panel Survey is collected by GfK NOP, the Office for National Statistics and the Northern Ireland Statistics and Research Agency under the management of the University of Essex's Institute for Social and Economic Research and funded by the Economic and Social Research Council. It is archived and distributed by the UK Data Archive. The original data creators, depositors and copyright holders, the funders of the Data Collections and the UK Data Archive bear no responsibility for their further analysis or interpretation.

The Expenditure and Food Survey was collected by the Office for National Statistics and sponsored by the Office for National Statistics and the Department for Environment, Food and Rural Affairs. It is distributed by the UK Data Archive, University of Essex, Colchester. The data is copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

9 REFERENCES

Engel, E. (1857) Die Produktions und Consumtionsverhaltnisse des Kaonigreichs Sachsen, reprinted with Engel (1895), Anlage 1, 1-54.

Engel, E. (1895) Das Lebenskosten belgischer Arbeiterfamilien frÄuher und jetzt, *Bulletin de Institut International de Statistique* 9: 1-124.

Expenditure and Food Survey (2006) Economic and Social Data Service. http://www.esds.ac.uk/findingData/snDescription.asp?sn=5986#doc

Holm, E., Mäkilä, K., and Lundevaller, E. (2009) 'Imitation or Interaction in Spatial Micro Simulation', Presentation at the 2nd General Conference of the International Microsimulation Association, Ontario, Canada. http://www.statcan.gc.ca/conferences/ima-aim2009/session3i-eng.htm#a44

Klevmarken, N. A., Andersson, I., Brose, P., Flood, L., Olovsson, P., and Tasiran, A. (1992) 'MICROHUS. A Micro-simulation Model for the Swedish Household Sector. A Progress Report'. Paper presented at the *International Symposium on Economic Modelling*, August 18-20, Gothenburg, Sweden.

Klevmarken, N.A. and P. Olovsson, (1996) "Direct and behavioural effects of income tax changes - simulations with the Swedish model MICROHUS", in A. Harding (ed.) *Micro-simulation and Public Policy*, Elsevier Science Publishers, Amsterdam

Klevmarken, N. A. (1997) 'Behavioural modelling in micro simulation models. A Survey', Papers 1997-31, Uppsala – Working Paper Series.

ISER (2008) University of Essex, Institute for Social and Economic Research, British Household Panel Survey: Waves 1-17, 1991 – 2008 [computer file]. 5th Edition. Colchester, Essex: UK Data Archive [distributor], March 2009. SN 5151.

Lawson, T. (2009) 'A Demographic Microsimulation Model Using NetLogo', Centre for Research in Economic Sociology and Innovation (CRESI), *Working Paper 2009-02*, Colchester: University of Essex.

ONS (2010) 'Household Final Consumption Expenditure Summary' 0.CN, *Consumer Trends – Q2 2010*, Office for National Statistics.

http://www.ons.gov.uk/ons/rel/consumer-trends/consumer-trends/q2-2010/index.html

