Modelling Business Energy Consumption using Agent-based Simulation Modelling

Jason Wong and Kay Cao

Analytical Services Branch, Australian Bureau of Statistics

Abstract

Simulation techniques are becoming more popular recently and have been applied to different research areas. An advantage of simulation over regression is that simulation based on micro-level data assumes heterogeneity between individuals. It also allows interaction between individuals and collective actions of agents impinge on the individual actions. This enables more sophisticated models to be built and overcomes the shortcomings of simplified ordinary least squares models. This paper developed a prototype agent-based simulation model to estimate business energy consumption taking into account business decision-making process in responding to government policies. The model developed in this study has been applied to synthetic data to assess its feasibility.

I. Introduction

Current modelling techniques employed for deriving non-survey energy consumption estimates often involved single equation approach. For example, Cao et al. (2013) used a log-linear model to estimate business energy consumption based on turnover and industry characteristics. This method required adjustment for log transformation bias and model misspecification to arrive at final estimates. Among the alternative modelling approaches suggested by the authors, agent based simulation method is one that could be considered given the availability of micro data.

Simulation techniques are becoming more popular recently and have been applied to many research areas. Agent based simulation assumes heterogeneity between individuals and allow

1 The authors would like to thank Ruel Abello and Anil Kumar (ABS Analytical Services) for their advice. Responsibility for any errors or omissions remains solely with the authors. The views in this paper are those of the authors and do not necessarily represent the views of the Australian Bureau of Statistics.
for their interactions. This enables more sophisticated models to be built and overcomes the shortcomings of over simplified Ordinary Least Squares models. Agent based simulation also allows for taking into account business decision making process in responding to government policies.

This paper attempts to develop a prototype agent based simulation model for business energy consumption, using data from the 2008-09 *Energy, Water and Environment Survey* (EWES) (ABS, 2010) and the *Business activity statement Unit Record Estimates* (BURE) data. Synthetic data, which were applied to the model, were generated based on the multivariate distribution of the required variables from these two surveys.

The remainder of the report is organised as follows. Section II describes Agent-based Simulation Modelling. Section III provides a literature review on the applications of Agent-based Simulation Modelling to the energy sector. Section IV presents a framework of the model for estimating business energy consumption. Section V discusses the data issues with regards to running the model described in Section IV. Section VI presents the results. Section VII concludes.

### II. Agent-based Simulation Modelling

As mentioned in the previous section, an advantage of Agent-based simulation over regression is that simulation based on micro-level data assumes heterogeneity between individuals. This is a reasonable assumption since in the case of energy consumption, it is impossible to accurately explain individual behaviour with available individual characteristics. Agent-based Simulation also allows interaction between individuals and collective actions of agents impinge on the individual actions of agents. Whereas in regression using micro-level data, each individual consumption is estimated under the assumption of homogeneity from which the aggregate consumption of the group is derived. Since each estimated value from a regression has a variance and is independent of each other, summing up the estimated values could lead to
significantly high variance in the estimate of interest. Further, Agent-based Simulation allows both cross-sectional micro-level (e.g. turnover of individual firm) data and time series macro-level (e.g. GDP) data to be included in the model. This enables more sophisticated models to be built and overcomes the shortcoming of over-simplicity in single equation modelling approach.

Although agent-based simulation modelling allows higher flexibility than regression in terms of the data and the assumptions of the model, one should not overlook the pitfalls. An agent-based simulation model is built upon the assumptions of the agents’ behaviour. It is imperative that the assumptions are correct. Any wrong assumptions would lead to invalid results. Unfortunately, information on agents’ behaviours is usually scarce and subjective assumptions are unavoidable. Moreover, simulation is computationally intensive and time-consuming. A balance between accuracy and computation time should be attained and overcomplicated models should be avoided.

A summary of the strengths and weaknesses of agent-based simulation modelling is presented in Appendix A.

A typical framework of an agent-based model consists of three elements (Macal et al., 2010):

- **Agents:** An agent can be an individual, an entity, an institution or government etc. Each agent has its own attributes and behaviour. An agent’s behaviour is driven by its goals (e.g. maximising profits while limiting level of risk). Rules, based on the behaviour of the agents, are set up to guide the actions of, and interactions between, agents. Agents are heterogeneous. Their actions depend on the information they have access to, the characteristics and resources of the agents, and their own models for decision making. This implies that agents who share the same characteristics and resources make different decisions. Agent interaction would also lead to different levels of resources accumulated by agents which cause further discrepancies in decision making of agents sharing the same characteristics.
• Relationship between agents: Which agents interact with which and how they interact are defined.

• Environment: Agents also interact with the environment. The environment is a set of conditions which all individuals face, e.g., GDP growth of the economy. Individuals obtain information from the environment to assist them in decision making.

An example of a feasible framework to model business energy consumption is provided in Section IV. It outlines the system of equations and illustrates the interactions between agents and the environment, as well as how agents’ decisions are made based on the outcomes of the previous period.

The simulation processes are essentially broken down into two parts. The first part estimates the parameters of the model, which may involve regression modelling or findings from other studies. The second part simulates the outputs of the models. Details on the simulation processes are described in Section 4.5.

III. Literature Review

The applications of Agent-based Simulation Modelling mainly fall in the disciplines of social science and biology. Recently, academics have applied Agent-based Simulation Modelling in the study of energy market. Wild et al. (2010) described the ANEM model, which is an application, of the Agent-based Simulation framework to the Australian National Electricity Market. The ANEM model is modified from the Agent-based Modelling of Electricity System (AMES) model developed in the USA (Sun et al., 2007) to account for the differences between the electricity markets in the two countries. There are two classes of agents involved in this model, namely Loading Serving Entities (LSE) and energy generators. LSE are obliged to provide electricity to residential, commercial and industrial end-users. They purchase electricity from the wholesale market and supply to end-users in the retail market. Generators produce electricity and sell in
the wholesale market. The transmission grid characteristics and hedging strategies of both LSE and generators are considered in the model. A system of equations is constructed and the simulated outcomes optimise the objective function while satisfying a set of constraints. The ANEM model was later used to assess the impact of carbon pricing on the electricity market in Australia in Wild et al. (2012). This paper evaluated the impact of carbon pricing on the electricity market under three perspectives: wholesale electricity prices; carbon-pass-through rates; and retail electricity tariffs. Simulated results were computed based on scenarios with different carbon prices. The results from different scenarios were compared to assess the impact of carbon pricing on the electricity market.

More relevant to this study, Martinez-Moyano et al. (2011) investigated the dynamics between willingness of adopting and choice of energy-efficiency technologies and energy consumption of commercial buildings. Factors such as building types, climate and energy-efficiency technology were considered. The willingness of adopting energy-efficiency technologies depends on the risk profiles of the building owners. Five scenarios were modelled. Scenario 1 assumes owners do not adopt any energy-efficiency technologies. Scenario 2 assumes owners randomly choose one technology should they decide to adopt any energy-efficiency technologies. In Scenario 3, owners choose the most effective four energy-efficiency technologies available to them if they decide to reduce energy consumption. Scenario 4 and 5 is analogous to Scenario 2 and 3, but with the intervention from the government to encourage adoption of energy-efficiency technologies. The simulated results are consistent with the prior expectation, with Scenario 1 having the poorest energy performance (measured in kWh/m²/year) and Scenario 5 the highest performance.

Zhou et al. (2007) provided a rich literature review on the implementation of Agent-based Simulation in various electricity markets.
Appendix B lists some fields other than energy where Agent-based Simulation Modelling has been applied to. These disciplines include, but not restricted to, ecological science, social science and economics.

**IV. An Example of Framework for Energy Modelling**

In this framework, we have two classes of agents, namely business and government. These agents interact with each other and impinge on the actions of others. The consequences of these actions would change the environment which would then have an impact on the actions of agents in the next period. This framework intends to show the concepts of Agent-based Simulation modelling and simplifies the real world. This framework can be expanded to have energy providers and other flows such as taxes, which would depend on data availability. Martinez-Moyano (2011) provides an example of modelling commercial buildings energy consumption using a similar but more sophisticated framework.

The following diagram summarises and explains the relationships.

4.1 Relationships between Agents and the Environment

![Diagram showing relationships between agents and the environment](image)

Note: 1, 2, 3 and 4 describe the interactions among agents and between agents and the environment:

2 The dotted simply means current consumption becomes last period consumption in the next simulation.
Environment’s Impact on Agents

1 After examining the expenses on energy from previous periods, businesses consider whether actions, such as undertaking energy-efficiency measures, are to be taken to reduce energy consumption in order to lower cost.

2 Government decides if any policies should be put in place to adjust energy consumption behaviour based on energy consumed in previous periods.

Agents’ Behaviour

3 Businesses’ decisions determine how much energy is consumed.

4 Government subsidies may be given to businesses to undertake energy-efficiency measures as a means to reduce total business energy consumption.

A system of equations is constructed based on these relationships. These equations could be a function of a number of variables or a distribution. The system of equations below is an example for estimating total energy consumption of a group of business firms.

4.1 Energy consumption of an individual firm:

\[ C_{i,t} = C_{i,t} \cdot \prod_{T=1}^{t} PEE_{i,T} \]  

(4.1)

\[ C_{i,t} = M_{i,Adj}.exp(\beta_1 + \beta_2 \ln(TO_{i,t}) + \beta_3 \ln(P_{i,t}) + \sum_{j=1}^{n-1} \gamma_j ind_j) \]  

(4.2)

where

\( C_{i,t} \) is the energy consumed by firm \( i \) at time \( t \), \( ind_i \) is the industry firm \( i \) belongs to;

\( PEE_{i,t} \) is the effect on energy consumption of energy-efficiency measures at time \( t \);
\( P_{i,t} \) is the price of energy paid by firm \( i \) at time \( t \);

\( TO_{i,t} \) is the turnover of firm \( i \) at time \( t \);

\( C_{i,t} \) is a latent variable which represents the energy consumed by firm \( i \) at time \( t \) without the influence of energy-efficiency measures.

\( M_i \) is estimated probability\(^3\) of unit \( i \) having non-zero energy consumption;

\( Adj \) is the smearing estimate\(^4\) of log transformation bias adjustment factor;

\( ind_j \) are industry dummies (here we categorise industries using the ANZSIC 2006 at two-digit level);

\( \beta \) and \( \gamma \) are coefficients from regression which are presented in Appendix C.

Based on (4.1) and (4.2), the energy consumed by a firm in a specific period depends on the industry it belongs to and the turnover of the period, as well as the cumulative energy efficiency investment.

4.2 Government subsidies for business uptake of energy-efficiency measures:

\[
GL_{t+1} = \begin{cases} 
1 & \text{if } \sum C_{i,t} / \sum C_{i,t-1} > c_1 \text{ for } t \geq 1 \\
0 & \text{Otherwise}
\end{cases}
\]

(4.3)

\(^3\) Estimated by logistic regression. Refer to Appendix D for the results.

\(^4\) \( Adj = \frac{\sum_{i=1}^{N} e^{u_i}}{N} \) where \( u_i \) is the regression residual and \( N \) is the regression sample size (Duan 1983). The estimated \( Adj \) is 2.5815.
\[ GT_{i,t+1} \sim U(d, e) \text{ if } GI_{iud,t+1} = 1 \]

Otherwise \( GT_{i,t+1} = 0 \) \hspace{1cm} (4.4)

where

\( GI \) is a binary variable of whether the government subsidises businesses to undertake energy-efficiency measures;

\[ \sum C \] is the total energy consumption of the industry.

\( GT \) represents the willingness of a firm to take up a risky investment in reducing energy consumption. The higher the ratio, the higher the proportion of energy expense needs to be reached before the owner invests in energy-efficiency measures which involve risk. Therefore, high ratio implies the owner is risk averse and vice versa;

\( c_i \) is a constant, describing the threshold above which the government would provide incentives for firm uptaking energy efficiency measures;

\( U \) denotes an uniform distribution\(^5\).

Equation (4.3) states if the total energy consumption increases by more than a certain percentage, the government will provide subsidies for energy-efficiency measures.

Equation (4.4) states should government subsidise businesses, its effect on risk profiles of individual firms to undertaking energy-efficiency measures is stochastic and follows a uniform distribution. The parameters \( d \) and \( e \) are negative since subsidies encourage businesses to invest in energy-efficiency measures.

\[^5\text{A distribution which all intervals of the same length are equally probable.}\]
4.3 Adoption of energy-efficiency measures from an individual firm:

\[ EE_{i,t+1} = 1 \text{ if } PC_{i,t} / TO_{i,t} > T_i + GT_{i,t}, \text{ where } T_i \sim N(x, y) \]
Otherwise \( EE_{i,t+1} = 0 \) \hspace{1cm} (4.5)

\[ PEE_{i,t+1} \sim U(a, b) \text{ if } EE_{i,t+1} = 1 \]
Otherwise \( PEE_{i,t+1} = 1 \) \hspace{1cm} (4.6)

\( EE \) is a binary variable representing whether a firm undertaking any energy-efficiency measures;

\( PEE \) represents the effect of the measures on energy consumption;

\( T \) is the initial (net of effect of government intervention) willingness of individual firms to undertake energy-efficiency;

\( N \) denotes a normal distribution.

In equation (4.5), it shows that when the expense on energy as a percentage of turnover in a period is higher than a threshold (i.e. the risk profile of the firm in this model), a firm will undertake some energy-efficiency measures.

Equation (4.6) states the impact of energy-efficiency measures on energy consumption is stochastic and follows a uniform distribution. The reasonable range of parameters \( a \) and \( b \) is \( 0 < a < b < 1 \). This implies that the energy-efficiency measures reduce energy consumption from previous year proportionally.

The chart below summarises the structure of the system of equations:
4.2 Flow Chart of Structure of Framework

As discussed in Section II, Agent-based Simulation Modelling is broken down into two parts. In Part 1, the parameters of the stochastic distributions, i.e. $x$, $y$, $a$, $b$, $c_1$, $d$, $e$ and the functional form and coefficients of $f_i$ are determined. The functional form of $f_i$ can be
determined from the literature and coefficients estimated by a regression. The remaining parameters are currently subjectively determined but could be improved with further subject matter knowledge.

4.5 Simulation Steps

The simulation of outputs is carried out once all the parameters are determined. In the first simulation \( (t = 0) \), it is assumed that there is no energy-efficiency measures taken by any business and no government subsidy and intervention in energy market. This assumption is made due to the lack of information to model \( GT_{i,1} \), \( PEE_{i,1} \) and \( A_t \) at this stage. To begin the simulation, \( TO_{i,0} \), \( ind_i \) and \( P_0 \) are inputted into the system of equations and the outputs for the period are obtained. Outputs from the previous simulation needed for the next period are the energy consumed \( C_{i,0} \), the impact of energy-efficiency measures by businesses \( PEE_{i,1} \). For the next period \( (t = 1) \), it is assumed that no government subsidy is given. \( TO_{i,1} \), \( ind_i \), \( P_{ind,1} \), \( PEE_{i,1} \) and \( C_{i,1} \) are inputted into the equations and outputs are obtained. From period \( (t = 2) \), the full model is run.
The chart below summarises the simulation processes:

4.3 Simulation Processes

V. Data Issues

It is of interest to put the framework described in the previous section in practice and apply some data to the model. The objective of the simulation in this paper is to estimate energy consumed in years (2008-09, 2009-10 and 2010-11) based on 2008-09 EWES and pseudo data for 2009-20 and 2010-11. Information relating to energy-efficiency measures, i.e. parameters in Equation (4.3), (4.4), (4.5) and (4.6), are not available at the moment, although such information may be obtained by further research in the literature or consulting the relevant experts in the field of study. For the purpose of demonstrating the use of Agent-based Simulation Modelling in estimating business energy consumption, assumptions of the parameters in Equation (4.3), (4.4), (4.5) and (4.6) have been made and presented in Appendix E.

Price data for 2008-09 were sourced from the EWES dataset. Producer price indexes were then used to derive unit level prices for 2009-10 and 2010-11. Turnover data were generated by using the pseudo dataset generation method as described in Appendix F.
VI. Estimation Results

The simulation is run for all 3 years of available data, namely 2008-09, 2009-10 and 2010-11. Following the simulation steps in Section 4.5, the results obtained are summarised in the table below:

Table 6.1: Simulated Total Energy Consumption of Manufacturing Industry

<table>
<thead>
<tr>
<th></th>
<th>2008-09</th>
<th>2009-10</th>
<th>2010-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Consumption (GJ)</td>
<td>$3.1782 \times 10^8$</td>
<td>$2.9556 \times 10^8$</td>
<td>$2.3612 \times 10^8$</td>
</tr>
</tbody>
</table>

Note that the above results were derived using the pseudo dataset for turnover. Given the lower turnover average (compared to those of 2008-09 BURE) and rising energy prices, the estimates are decreasing over the years.

VII. Conclusion

This report presents an alternative method to modelling business energy consumption with the use of Agent-based Simulation Modelling. Issues regarding the implementation of the model are discussed and a solution to the data issues is proposed and carried out in this study.

The aim of this project is to explore the possibility of using Agent-based Simulation Modelling to model business energy consumption rather than obtaining reliable estimates of energy consumption, therefore, some parameters are subjectively chosen. Digging into the literature and available data sources in order to refine the parameters could be an area for future research.

Replacing estimated turnover data with real turnover data may help to improve model estimates. This will be considered in the next phase of this research project.
References


Appendix

A. Strengths and Weaknesses of Agent-based Simulation Modelling

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allow for heterogeneity between individuals</td>
<td>Wrong assumptions invalidate results</td>
</tr>
<tr>
<td>Allow interaction between individuals</td>
<td>Subjective assumptions are unavoidable and may be wrong</td>
</tr>
<tr>
<td>Collective actions of each agent impinge on individual actions of agents</td>
<td>Computational intensive</td>
</tr>
<tr>
<td>Flexibility in using data from various data sources</td>
<td></td>
</tr>
</tbody>
</table>

B. Applications of Agent-based Simulation Modelling

<table>
<thead>
<tr>
<th>Area</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Science: Spread of contagious disease</td>
<td>Eubank et al. (2004)</td>
</tr>
<tr>
<td>Social Science: Shopping time in retail store</td>
<td>North et al. (2007)</td>
</tr>
<tr>
<td>Social Science: Behaviour in emergency evacuation</td>
<td>Pan et al. (2007)</td>
</tr>
<tr>
<td>Economics: Innovation diffusion</td>
<td>Garcia et al. (2011)</td>
</tr>
<tr>
<td>Economics: Labour market</td>
<td>Leombruni et al. (2006)</td>
</tr>
</tbody>
</table>
### C. OLS for Energy Consumption

Dependent Variable: Log of Total Energy Consumption (GJ)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-6.4029</td>
<td>0.2582</td>
</tr>
<tr>
<td>Log turnover</td>
<td>0.6627</td>
<td>0.0159</td>
</tr>
<tr>
<td>Log Energy Price</td>
<td>-1.1605</td>
<td>0.0313</td>
</tr>
<tr>
<td>ind11 (=reference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ind12</td>
<td>-0.1194</td>
<td>0.1709</td>
</tr>
<tr>
<td>ind13</td>
<td>-0.7381</td>
<td>0.1283</td>
</tr>
<tr>
<td>ind14</td>
<td>0.2425</td>
<td>0.1652</td>
</tr>
<tr>
<td>ind15</td>
<td>-0.3901</td>
<td>0.1786</td>
</tr>
<tr>
<td>ind16</td>
<td>-0.4567</td>
<td>0.1999</td>
</tr>
<tr>
<td>ind17</td>
<td>-0.3148</td>
<td>0.2064</td>
</tr>
<tr>
<td>ind18</td>
<td>-0.4095</td>
<td>0.1366</td>
</tr>
<tr>
<td>ind19</td>
<td>-0.2234</td>
<td>0.1719</td>
</tr>
<tr>
<td>ind20</td>
<td>-0.0855</td>
<td>0.1353</td>
</tr>
<tr>
<td>ind21</td>
<td>-0.2910</td>
<td>0.1424</td>
</tr>
<tr>
<td>ind22</td>
<td>-0.4798</td>
<td>0.1234</td>
</tr>
<tr>
<td>ind23</td>
<td>-0.9065</td>
<td>0.1375</td>
</tr>
<tr>
<td>ind24</td>
<td>-0.6714</td>
<td>0.1248</td>
</tr>
<tr>
<td>ind25</td>
<td>-0.6153</td>
<td>0.1552</td>
</tr>
</tbody>
</table>

Adjusted R² 0.7068
RMSE 1.2326
Observations 1573

*, ** and *** represents the coefficient is statistically significant at 10%, 5% and 1% level, respectively.
D. Logistic for Non-zero Energy Consumption

Dependent Variable: Non-zero Energy Consumption (GJ)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-3.0635 ***</td>
<td>0.668</td>
</tr>
<tr>
<td>Log turnover</td>
<td>0.3724 ***</td>
<td>0.047</td>
</tr>
<tr>
<td>ind11 (=reference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ind12</td>
<td>0.6008</td>
<td>0.5568</td>
</tr>
<tr>
<td>ind13</td>
<td>-0.0876</td>
<td>0.3237</td>
</tr>
<tr>
<td>ind14</td>
<td>0.3069</td>
<td>0.5065</td>
</tr>
<tr>
<td>ind15</td>
<td>0.4578</td>
<td>0.6309</td>
</tr>
<tr>
<td>ind16</td>
<td>0.4199</td>
<td>0.6312</td>
</tr>
<tr>
<td>ind17</td>
<td>1.6279</td>
<td>1.056</td>
</tr>
<tr>
<td>ind18</td>
<td>-0.2021</td>
<td>0.3751</td>
</tr>
<tr>
<td>ind19</td>
<td>0.6965</td>
<td>0.6294</td>
</tr>
<tr>
<td>ind20</td>
<td>1.0018 *</td>
<td>0.5458</td>
</tr>
<tr>
<td>ind21</td>
<td>-0.08</td>
<td>0.3874</td>
</tr>
<tr>
<td>ind22</td>
<td>0.928 **</td>
<td>0.4605</td>
</tr>
<tr>
<td>ind23</td>
<td>-0.1257</td>
<td>0.3765</td>
</tr>
<tr>
<td>ind24</td>
<td>0.3464</td>
<td>0.3833</td>
</tr>
<tr>
<td>ind25</td>
<td>0.9262 *</td>
<td>0.5498</td>
</tr>
</tbody>
</table>

Adjusted R² 0.1188

Per-cent Concordance 71.6%

Observations 1700

*, ** and *** represents the coefficient is statistically significant at 10%, 5% and 1% level, respectively.
E. Parameters Used in Simulation Framework

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x) (^6)</td>
<td>0.00026</td>
</tr>
<tr>
<td>(y) (^7)</td>
<td>0.000065</td>
</tr>
<tr>
<td>(a)</td>
<td>0.75</td>
</tr>
<tr>
<td>(b)</td>
<td>0.95</td>
</tr>
<tr>
<td>(c_1)</td>
<td>1.3</td>
</tr>
<tr>
<td>(d)</td>
<td>-0.0002</td>
</tr>
<tr>
<td>(e)</td>
<td>0</td>
</tr>
</tbody>
</table>

F. Pseudo Dataset Generation

A dataset which replicates the 2008-09, 2009-10 and 2010-11 BURE dataset using the bivariate distribution is generated by the following procedures:

- 2008-09 BURE dataset is replicated\(^8\) using the bivariate distribution of log of turnover and log of derived energy price. This process is done separately for each industry at the two-digit ANZSIC 2006 level. Assuming log of turnover and log of derived energy price are both normally distributed, each bivariate distribution is defined by the normal distributions of the two variables and the correlation between the two.

- Assuming individual firms’ energy prices inflate consistently with the Producer Price Index for electricity input and the correlation between log of turnover and log of

\(^6\) Mean is chosen as the 90\(^{th}\) percentile of \(P_{C_{i,t}}/TO_{i,t}\) from the 2008-09 data, so that it is expected that 10% of the firms would take up an energy-efficiency measures when there are no subsidies.

\(^7\) The variance is chosen, arbitrarily, as one-fourth of the mean.

\(^8\) We only generate turnover and derived energy price.
derived energy price in 2008-09 is also valid for subsequent years, log of turnover for subsequent years, \( \ln TO_{i,t} \), are estimated using the following formulae:

\[
E(\ln TO_{i,t} \mid \ln P_{i,t}) = \mu_2 + \rho \sigma_2 (\ln P_{i,t} - \mu_1) / \sigma_1
\]  
(6.2)

\[
\ln TO_{i,t} = E(\ln TO_{i,t} \mid \ln P_{i,t}) + \sigma_3
\]  
(6.3)

where

\( \ln TO_{i,t} \) is the log of turnover of firm \( i \) at time \( t \);

\( \ln P_{i,t} \) is the log of derived energy price of firm \( i \) at time \( t \);

\( \rho \) is the correlation between \( \ln TO_{i,t} \) and \( \ln P_{i,t} \) based on 2008-09 data;

\( \ln TO_{i,t} \sim N(\mu_2, \sigma_2^2) \) based on data at time \( t \);

\( \sigma_1 \) is the standard deviation of log of derived energy price based on 2008-09 data;

\( \sigma_3 = \sigma_2 \sqrt{1 - \rho^2} \) is the standard deviation of \( E(\ln TO_{i,t} \mid \ln P_{i,t}) \);

\( \mu_1 = \mu_{i,08-09} + \ln(r) \)

where

\( \mu_{i,08-09} \) is the mean of log of derived energy price based on 2008-09 data;

\( r \) is the rate of change in Producer Price Index for electricity input between 2008-09 and time \( t \).
Exits and new entries into BURE are also taken into account. Each year, a number of firms, equal to the number which left the BURE dataset, are dropped from the pseudo dataset. Similarly, a number of firms are added to the pseudo dataset, with the bivariate distribution based on the BURE data of the corresponding year.