**How can micro data registers improve the estimates of real estate in the national accounts?**

Christian Gysting, cgy@dst.dk  
Gitte Frej Knudsen, gfk@dst.dk  
*Statistics Denmark*

**Abstract:** A variety of registers at Statistics Denmark makes it possible to calculate the real estate at market prices on an individual level for households and companies. The market price is calculated from actual sales of real estate on the market.

The paper will discuss how the valuation of the real estate in the national accounts can benefit from the estimation of a market price on macro level generated from the micro register data with real estate. The paper focus on dwellings (buildings), but the suggested methods could be used for non-residential buildings as well.

The conclusion is that using a micro based data either in combination with macro data or complete use of micro data in a hedonic set up should make it possible to calculate a complete set of balance sheets – separate but consistent values for dwellings and land – based on real estate data at market prices in the national accounts.

1. **Households’ real estate on individual level**

1.1 **Results**

*Register with market value*  
Statistics Denmark has established an individual-based register with the market value of the households' real estate. The register includes information about the market value on both property level and individual level, as the market value of each individual's part of a property has been calculated. An individual in the register is a person or a sole proprietorship, which is registered with a personal ID number or a business register number. This means that households are defined as the household sector in the national accounts.

*Aggregated micro data*  

<table>
<thead>
<tr>
<th>Billion DKK (current prices)</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owner occupied dwellings</td>
<td>2.446</td>
<td>2.837</td>
<td>3.107</td>
<td>3.913</td>
<td>4.005</td>
<td>3.534</td>
<td>3.476</td>
</tr>
<tr>
<td>Co-operative dwellings 1</td>
<td>134</td>
<td>152</td>
<td>253</td>
<td>228</td>
<td>244</td>
<td>228</td>
<td>228</td>
</tr>
<tr>
<td>Real estate, % of GDP</td>
<td>176</td>
<td>193</td>
<td>206</td>
<td>244</td>
<td>242</td>
<td>226</td>
<td>211</td>
</tr>
</tbody>
</table>

*Table 1: Households’ real estate (aggregated micro data)*

(Housing co-operative is a legal entity - usually cooperation - that owns residential buildings.

The households’ real estate is estimated to 3,704 billion DKK in 2010, of which 6 per cent of the value is co-operative dwellings. The market value of the households' real estate is largest in 2007 and 2008, where the economic expansion had the maximum effect on the real estate market in Denmark. It is possible to extend the register to include real estate owned by others than households (e.g. companies).

1 Housing co-operative is a legal entity - usually cooperation - that owns residential buildings. The co-operative is membership based, with membership granted by a share purchase in the co-operative. Each shareholder is granted the right to occupy one housing unit.)
1.2 Market value of owner occupied dwellings

Official real estate valuations

The official real estate valuations from the Danish tax authorities do not reflect the market prices. The difference between the official real estate valuations and the market price depends on the market conditions, which are not described to a sufficient degree in the official real estate valuations.

The adjustment factor

By linking the actual real estate sales with the official real estate valuations, it is possible to calculate the average relationship between the actual price and the real estate valuation for a dwelling. This relationship is called the “Adjustment factor”. In the model it is assumed that the adjustment factor between the actual purchase price (the market value) and the official real estate valuation is the same for individual types of real estate (e.g. one-family houses) within the same geographical area (e.g. a postal code). The market values of the dwellings that have not been sold are calculated by multiplying the adjustment factor with the official real estate valuation.

Table 2: Market value by category of property (aggregate micro data)

<table>
<thead>
<tr>
<th>Billion DKK (current prices)</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-family houses</td>
<td>1.412</td>
<td>1.632</td>
<td>1.960</td>
<td>2.273</td>
<td>2.252</td>
<td>1.945</td>
<td>1.989</td>
</tr>
<tr>
<td>Dwelling houses</td>
<td>131</td>
<td>149</td>
<td>148</td>
<td>220</td>
<td>226</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Flats (in a block of flats)</td>
<td>198</td>
<td>249</td>
<td>310</td>
<td>333</td>
<td>324</td>
<td>275</td>
<td>293</td>
</tr>
<tr>
<td>Holiday cottages</td>
<td>173</td>
<td>215</td>
<td>235</td>
<td>288</td>
<td>288</td>
<td>252</td>
<td>256</td>
</tr>
<tr>
<td>Farms</td>
<td>351</td>
<td>391</td>
<td>255</td>
<td>528</td>
<td>633</td>
<td>598</td>
<td>490</td>
</tr>
<tr>
<td>Other</td>
<td>180</td>
<td>201</td>
<td>200</td>
<td>271</td>
<td>283</td>
<td>265</td>
<td>249</td>
</tr>
<tr>
<td>Total</td>
<td>2.446</td>
<td>2.837</td>
<td>3.107</td>
<td>3.913</td>
<td>4.005</td>
<td>3.534</td>
<td>3.476</td>
</tr>
</tbody>
</table>

Other use of statistics on individual level

Registers on individual level can be used for distribution analyses, e.g. in relation to age, income, family type or socioeconomic status.

1.3. Valuation of owner-occupied dwellings in the national accounts

Real estate in the balance sheets

In the balance sheets of the national accounts the value of real estate (including structures and land improvement) can be calculated by adding several items; Dwellings (AN111), other buildings and structures (AN112) and land (AN211). All items can be sub-divided into type of owner by institutional sector (i.e. non-financial sector (S11) or households (S14)). Further, dwellings can be sub-divided into the type of user of the dwellings; Owner-occupied dwellings and rented non-residential buildings.

Calculations in practice

In practice, most of the items can’t be separately identified at market prices. Either is it necessary to use alternative methods for estimating the values, for dwellings and other buildings and structures the most common approach is the perpetual inventory method (PIM), or by calculating the combined value of land and buildings and somehow separate the value into its land component and its building component.

PIM approach

Today Statistics Denmark only publishes the value of dwellings and other buildings and structures. The value of land is not published. Statistics Denmark uses the PIM approach to the calculation.

2 Statistic Denmark is working on a project regarding household wealth. The valuation is based on registers specified on individual level with the market value of real estate and cars as well as financial assets and liabilities (including employment-related pension schemes).
The national accounts estimate the value of buildings at reconstruction prices. There are two concepts for real estate, which are “Gross stock” and “Net stock”. Gross stock is the value of buildings at the reconstruction price, which leave out life span and wear in the calculation. The calculation of the net stock adjusts for all depreciations and therefore the net stock is an estimate of the market price.

The individual-based register with the estimate of the households’ real estate are valuated with a total market price of buildings and land. The national accounts estimates the value of buildings based on macro data, and therefore the two statistics are not directly comparable. The land component could be estimated as the difference between the value of the aggregated micro data (Household Wealth Project) and the macro estimation of buildings (national accounts). See figure 1.

**Figure 1: Market value by national accounts and aggregated micro data**

The large price increases will implicitly be a part of the land value. The market prices calculated on basis of the actual real estate sales in the period gives an image of the price bubble that was present in the real estate market. The national accounts’ calculation of the value of buildings is based on the construction cost index for buildings and will as such not reflect the economic trends in the real estate market to the same degree.

**2. New method to valuation of the real estate in the national accounts**

If the results are used from the real estate wealth project to measure the total value of real estate for dwellings, additional estimations is necessary for compiling values for land and/or buildings (dwellings). This paper suggests 2 different approaches:

1. **The residual approach.** The micro based data is used to estimate the combined value of land and buildings. And with the combined value and the value of buildings available (from PIM estimations), the land component could be estimated as a residual.
(2) **The hedonic approach.** The micro based data could be used to estimate the average value for one square meter of land and one square meter of building. The estimated parameters can be used to compile the values for land at macro level.

In the next two chapters the two approaches will be described with respect to data requirements and method. Both approaches make directly or indirectly use of micro data registers.

### 2.1 The residual approach

The residual approach is based upon a very simple equation; the total value of real estate (T) is equal to the value of land (L) and the value of buildings (B). If this approach is applied, it is normal to calculate the values for buildings and the total value for real estate, and then estimate the value of land as a residual.

Most national statistical offices already compile values for buildings, most commonly by using the Perpetual Inventory Method (PIM). Requirements for compiling values for buildings by PIM is assumptions on service lives, depreciation profile and discard pattern together with time series of gross fixed capital formations for buildings (dwellings). The Perpetual Inventory Method is a macro level calculation based on aggregate values for gross fixed capital formations.

The second requirement for applying the residual approach is values for real estate for dwellings. Chapter 1.1 and 1.2 describe how these numbers can be compiled at micro level by using register data. With values available for real estate and buildings, figures for land can easily be compiled by subtracting the value of buildings from the value of the real estate.

The residual approach combines a macro level calculation of the value of buildings (PIM) with a micro level calculation of the value of real estate. The residual approach must be assumed to produce reliable results as long as the market for real estate is not subject to large negative impacts.

### 2.2 The hedonic approach

The hedonic approach is a conceptually different approach then the residual approach. It is based on micro data and makes use of a regression model. Required source data is actual sales prices for real estates and the corresponding numbers of square meter of buildings and land. Information on age of the buildings would properly be a useful additional input into the calculations. All these data is available at Statistics Denmark.

The hedonic regression model can in its most simplistic form be written as:

\[ Y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i \quad i = 1, \ldots, n. \]

---

The outputs of the model are values for 2 parameters; \( \beta_1 \), which is the price per square meter of building, and \( \beta_2 \), which is the price for one square meter of land. Input to the model is the dependent variable \( Y_i \), which is the real estate price for observation number \( i \), and 2 independent variables; \( x_{1i} \), which is the size of the building measured in square meters for observation number \( i \), and \( x_{2i} \), which is the size of the land measured in square meters for observation number \( i \). The sample contains of \( n \) observations, and \( \epsilon_i \) is the error term.

**Interpretation of estimated parameters**

The parameters \( \beta_1 \) and \( \beta_2 \) are the central output of the hedonic regression model, they represent the price per square meter of building and land. Estimates for the total value of land can be obtained by multiplying price per square meter of land with the total number of square meter of land under buildings (dwellings).

**Location matters**

The estimated price per square meter of land depends on the size and boundary of the sample. If all observations for one country are merged into one group, it is unlikely that the results would be reliable. It is more reasonable to sub-divide the observations into groups with similar characteristics, most important with respect to location. The price of land normally depends heavily on location. Subdividing of data into sub-groups with different location is possible by using micro data from Statistics Denmark’s registers with real estate data.

**Use of hedonics in practice**

The use of hedonics for measuring value of land is limited. The method has been pioneered by Diewert and van der Haan and Hendriks (2010) which has used the method for sub-dividing the a residential property price index into 2 separate indices, one index for land and one for buildings. Further, using the method is very demanding because it requires technical skills and is work and data intensive.

### 3. Conclusion

Statistics Denmark has access to information on square meters of land and square meters of buildings, tax valuation of buildings and transaction prices for all traded real property in Denmark. Combined with national accounts figures for gross fixed capital formation and capital stocks for buildings (dwellings), this set of micro data makes it possible to compile balance sheets for Denmark which is in line with observed market prices. This article outlines 2 different approaches for compiling the currently missing part – the value of land.

The first approach – the residual approach – is the easiest approach to implement in practise. If the residual approach is selected, focus should be on compiling combined values for land and buildings for all types of of buildings, including non-residential buildings and rented residential buildings. The technical requirements are moderate but the approach is vulnerable to large price drops on real estate market.

The second approach – the hedonic approach – is technically difficult to implement in practice and work intensive. The approach requires estimation of a large number of regression models, including checking of the results.
However, the results must be assumed to be reliable also during periods with large price drops.

Because the hedonic approach is technically difficult and work intensive it is the opinion of the authors that the hedonic approach should only be used as a last resort if the residual approach does not result in reliable results.

4. References
