Supplement to "Convenient links for the estimation of hedonic price indexes: the case of unique, infrequently traded assets"

Esmeralda A. Ramalho and Joaquim J.S. Ramalho

Department of Economics and *CEFAGE-UE, Universidade de Évora*

November 2013

1 Introduction

This Supplement to the paper "Convenient links for the estimation of hedonic price indexes: the case of unique, infrequently traded assets" expands on the material contained there in Section 3.2. In particular, the analysis undertaken for arithmetic price indexes in Section 3.1 of the paper is here replicated to the case of geometric indexes.

2 Link between the price index formula and the form of the dependent variable of the hedonic function in the framework of geometric indexes

2.1 The case of a log-linear hedonic function

In the case of geometric price indexes, it is the use of a log-linear hedonic function that simplifies considerably the construction of quality-adjusted indexes. Indeed, when the data generating process of dwelling prices is suitably described by a log-linear model, a consistent predictor of the logged price is given by $\widehat{\ln(p_{it})} = x_{it}\hat{\beta}_t$ and, therefore, a consistent estimator for I_s^G is simply

$$\hat{I}_s^G = \frac{\exp\left(\frac{1}{N_s}\sum_{i=1}^{N_s} x_{is}\hat{\beta}_s\right)}{\exp\left(\frac{1}{N_0}\sum_{i=1}^{N_0} x_{i0}\hat{\beta}_0\right)}, \tag{1}$$

which can be decomposed into a quality ($\hat{I}_s^{G_q}$) and a quality-adjusted ($\hat{I}_s^{G_p}$) price index:

$$\hat{I}_s^G = \underbrace{\frac{\exp\left(\frac{1}{N_s}\sum_{i=1}^{N_s} x_{is}\hat{\beta}_b\right)}{\exp\left(\frac{1}{N_0}\sum_{i=1}^{N_0} x_{i0}\hat{\beta}_b\right)}}_{\hat{I}_s^{G_q}} \underbrace{\frac{\exp\left(\frac{1}{N_a}\sum_{i=1}^{N_a} x_{ia}\hat{\beta}_s\right)}{\exp\left(\frac{1}{N_a}\sum_{i=1}^{N_a} x_{ia}\hat{\beta}_0\right)}}_{\hat{I}_s^{G_p}}, \tag{2}$$

where $\hat{I}_s^{G_q}$ and $\hat{I}_s^{G_p}$ are consistent estimators for $I_s^{G_q}$ and $I_s^{G_p}$.

2.2 The case of an exponential hedonic function

On the other hand, when the true hedonic function has an exponential form, specification and estimation of an exponential regression model yields directly consistent estimates for the dwelling price, $\hat{p}_{it} = \exp\left(x_{it}\hat{\beta}_t^*\right)$, not for the logged prices that appear in the geometric price index formula. Moreover, the naive estimator given by the logarithm of $\hat{p}_{it}$, $\ln(\hat{p}_{it}) = x_{it}\hat{\beta}_t^*$, is not in general a consistent estimator for $E[\ln(p_t)|x_{it}]$. Indeed, it follows that

$$E[\ln(p_{it})|x_{it}] = x_{it}\beta_t^* + E(u_{it}^*|x_{it}), \tag{3}$$

where $E(u_{it}^*|x_{it}) \neq 0$.

To the best of our knowledge, the problem of going from level predictions to log predictions has never been analyzed in the econometrics literature. However, this is a very similar issue to that created by the assumption of a log-linear hedonic function in the arithmetic framework, being necessary to estimate a bias correction. Let

$$E(u_{it}^*|x_{it}) \equiv \mu_{it}^* = h(x_{it}^*\alpha_t^*), \tag{4}$$

where $h(\cdot)$ may be a nonlinear function and α_t^* is a vector of parameters. Assuming, for the moment, that $h(\cdot)$ is known and that a consistent estimator for α_t^*, $\hat{\alpha}_t^*$, is available, then, a consistent estimator for $\ln(p_{it})$ is given by $\widehat{\ln(p_{it})} = x_{it}\hat{\beta}_t^* + \hat{\mu}_{it}^*$, which yields the following estimator for the unadjusted geometric price index:

$$\hat{I}_s^G = \frac{\exp\left\{\frac{1}{N_s}\sum_{i=1}^{N_s}\left[x_{is}\hat{\beta}_s^* + h(x_{is}^*\hat{\alpha}_s^*)\right]\right\}}{\exp\left\{\frac{1}{N_0}\sum_{i=1}^{N_0}\left[x_{i0}\hat{\beta}_0^* + h(x_{i0}^*\hat{\alpha}_0^*)\right]\right\}}. \tag{5}$$

This estimator may be decomposed as follows:

$$\hat{I}_s^G = \underbrace{\frac{\exp\left\{\frac{1}{N_s}\sum_{i=1}^{N_s}\left[x_{is}\hat{\beta}_b^* + h\left(x_{is}^*\hat{\alpha}_b^*\right)\right]\right\}}{\exp\left\{\frac{1}{N_0}\sum_{i=1}^{N_0}\left[x_{i0}\hat{\beta}_b^* + h\left(x_{i0}^*\hat{\alpha}_b^*\right)\right]\right\}}}_{\hat{I}_s^{Gq}} \underbrace{\frac{\exp\left\{\frac{1}{N_a}\sum_{i=1}^{N_a}\left[x_{ia}\hat{\beta}_s^* + h\left(x_{ia}^*\hat{\alpha}_s^*\right)\right]\right\}}{\exp\left\{\frac{1}{N_a}\sum_{i=1}^{N_a}\left[x_{ia}\hat{\beta}_0^* + h\left(x_{ia}^*\hat{\alpha}_0^*\right)\right]\right\}}}_{\hat{I}_s^{Gp}}. \tag{6}$$

Therefore, unless $\mu_{it}^* = \mu^*$, the estimation of quality-adjusted geometric price indexes based on exponential hedonic functions requires the previous estimation of $\hat{\alpha}_s^*$ and $\hat{\alpha}_0^*$.

Similarly to the arithmetic case, some assumptions may be made in order to simplify the estimation of μ_{it}^*. In particular, we may compute the smearing-type estimator

$$\hat{\mu}_t^* = \sum_{i=1}^{N_t} \hat{u}_{it}^*, \tag{7}$$

provided that $E\left(u_{it}^*|x_{it}\right)$ does not depend on x_{it}; or we may assume that $\exp\left(u_{it}^*|x_{it}\right)$, in addition to unit mean, has a lognormal distribution with variance $\left[\exp\left(x_{it}^*\alpha_t^*\right) - 1\right]$ such that u_{it}^* has a normal (conditional) distribution with mean given by

$$\mu_{it}^* = -0.5 x_{it}^*\hat{\alpha}_t^*, \tag{8}$$

where $\hat{\alpha}_t^*$ results from regressing the squared residuals of the exponential hedonic function (plus one) on $\exp\left(x_{it}^*\alpha_t^*\right)$.

2.3 The general case

Consider now the estimation of a log-linear model when the true hedonic function has an exponential form and it is further assumed that $h(\cdot)$ is a linear function:

$$\mu_{it}^* = x_{it}^*\alpha_t^*. \tag{9}$$

Then,

$$E\left[\ln\left(p_{it}\right)|x_{it}\right] = x_{it}\beta_t^* + x_{it}^*\alpha_t^* = z_{it}\delta_t, \tag{10}$$

which shows that assumptions for μ_{it}^* of the type made in (9), such as those underlying the smearing and the normal-smearing estimators, are easily accommodated by the log-linear model.

Irrespective of assumption (9) being true or not and irrespective of the true generating process of dwelling prices, using a log-linear hedonic function is the only form of ensuring that no bias corrections are necessary for computing GHPI. Hence:

Link 1b: *There exists a link between the computation of GHPI and hedonic functions that consider logged asset prices as dependent variable.*