Appendix A: estimation method

In order for the Heckman–type selection model to be identified, the product innovation equation (the "selection equation") needs to contain variables that are not part of the product innovation advertising equation. These are the so–called "exclusion restrictions" that are summarized by the vector $\boldsymbol{z_i}$, where the subscript i denotes the ith observation. These exclusion restrictions must be orthogonal to the product innovation advertising decision.
In addition to the exclusion restrictions, the model for product innovation contains a set of variables that appear in both the product innovation and the product innovation advertising equation, vector $\boldsymbol{x_i}$.
Apart from this vector of joint variables $\boldsymbol{x_i}$, the product innovation advertising equation must also consist of a vector of variables $\boldsymbol{w_i}$ that appears in the advertising equation only. These again are exclusion restrictions, this time variables that affect product innovation advertising but not product innovation.
The *structural* equations for latent product innovation, PI_i^*, (as in all binary choice model the econometrician only observes the binary outcome but not the actual propensity of product innovation) and latent product innovation advertising, PIA_i^* respectively are hence:

$$\begin{aligned} \mathrm{PI}_i^* &= \delta\mathrm{PIA}_i^* + \boldsymbol{\alpha z_i} + \boldsymbol{\gamma}_{PI}\boldsymbol{x_i} + \mu_i && (1) \\ \mathrm{PIA}_i^* &= \boldsymbol{\gamma}_{PIA}\boldsymbol{x_i} + \boldsymbol{\theta w_i} + \eta_i, && (2) \end{aligned}$$

where PIA_i^* is only observed if product innovation took place, e.g. $\mathrm{PI}_i^* > 0$ in this binary choice setting. The terms μ_i and η_i denote are assumed to be i.i.d. bivariate normal distributed. The correlation between the two error terms is denoted by ρ.
Equation (3) cannot be directly be estimated because product innovation advertising is only observed if product innovation took place so that the information on product innovation advertising is a perfect predictor for product innovation (apart from the fact that it is also potentially endogenous). Likewise, Equation (2) cannot be directly estimated because it is conditional on product innovation. I solve this problem by proceeding in a standard way and estimate Equation (3) and (2) in *reduced form*. The estimation equation for product innovation hence is:

$$\mathrm{PI}_i^* = \delta(\boldsymbol{\gamma}_{PIA}\boldsymbol{x_i} + \boldsymbol{\theta w_i}) + \boldsymbol{\alpha z_i} + \boldsymbol{\gamma}_{PI}\boldsymbol{x_i} + \mu_i. \quad (3)$$

Equation (3) can consistently and, if the $\rho = 0$, efficiently estimated by a simple binary probit model. Apart from the possibility that there might be correlation between the error terms ($\rho \neq 0$), the interest is on the product innovation advertising as well so I estimate the equation for product innovation advertising, Equation (2), joint with the reduced form product innovation equation, Equation (3) using a Heckman–type model for two binary outcomes.

The drawback of such a reduced form estimation of the product innovation equation is that the effect of product innovation advertising on product innovation, the parameter δ, is not identified. In a second step following the estimation of the Heckman–type model I substitute these fitted values for latent product innovation advertising from the product innovation advertising equation in the Heckman–type model, $\mathrm{PIA}_i^* = \hat{\boldsymbol{\gamma}}_{PIA}\boldsymbol{x_i} + \hat{\boldsymbol{\theta}}\boldsymbol{w_i}$, into the structural form equation for product innovation, Equation (3).