Stockholms universitet

Research Report

Department of Statistics

No. 2018:6

A comparison of Stratified simple random sampling and Probability proportional-to-size sampling

Edgar Bueno

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

A comparison of Stratified simple random sampling and Probability proportional-to-size sampling

Edgar Bueno

Abstract

The sampling strategy that couples probability proportional-to-size sampling with the GREG estimator has sometimes been called "optimal", as it minimizes the anticipated variance. This optimality, however, relies on the assumption that the finite population of interest can be seen as a realization of a superpopulation model that is known to the statistician. Making use of the same model, the strategy that couples model-based stratification with the GREG estimator is an alternative that, although theoretically less efficient, has shown to be sometimes more efficient than the so-called optimal from an empirical point of view. We compare the two strategies from both analytical and simulation standpoints and show that optimality is not robust towards misspecifications of the model. In fact gross errors may be observed when a misspecified model is used.

Keywords: Survey sampling; Optimal strategy; GREG estimator; Model-based stratified sampling; Probability propotional-to-size sampling.

1 Introduction

When planning the *sampling strategy* (i.e. the couple *sampling design* and *estimator*) in a finite population survey setup, the statistician is often looking for "the most" efficient strategy. Godambe (1955), Lanke (1973) and Cassel et al. (1977) show that there is no uniformly best estimator, in the sense of being best for all populations. There is no best design either. Nevertheless, it is often possible to identify a set of strategies that can be considered as candidates. Our task is to choose one among this set. The "industry standard" for busines surveys, for example, has since long been stratified simple random sampling. The population is stratified into industry and within industry by some size variable. An alternative design, which is also often used, is probability proportional-to-size sampling.

The setup that will be used throughout this paper is as follows. We are interested in the estimation of the total of a study variable. The values of an auxiliary variable are known from the planning stage for all the elements. We will assume that ideal survey conditions hold. The auxiliary variable can be used at the design stage, the estimation stage or both, for obtaining an efficient strategy, where efficiency will be understood in terms of design-based variance.

The strategy that couples proportional-to-size sampling with the regression estimator (denoted πps–reg) has sometimes been called optimal (see, for example, Särndal et al. (1992), Brewer (1963), Isaki and Fuller (1982)). This optimality, however, relies on a superpopulation model which might not (and most certainly will not) hold exactly in practice. Wright (1983) proposed strong model-based stratification, which, making use of the same superpopulation model, defines a sampling strategy that couples stratified simple random sampling with the regression estimator.

Both strategies mentioned above rely on the assumption that the finite population can be seen as a realization of a particular model (section 2.2). The aim of this paper is to compare these strategies and try to answer the following question: is πps–reg still the best strategy when the model is misspecified? Besides the two strategies already mentioned, three more will be included in the study.

There are articles focused on a particular concrete situation, for example Kozak and Wieczorkowski (2005) who study πps and stratified designs in an agricultural survey. Rosén (2000a) investigates optimality of πps by means of simulations and theory. Holmberg and Swensson (2001) present a minor simulation study comparing these strategies. Our intention is to compare them from both analytical and simulation standpoints.

The contents of the article are arranged as follows. The framework is defined in section 2, where the estimators and designs of interest, as well as the superpopulation model, are presented. In section 3 we verify empirically the optimality of πps–reg under a correctly specified model. The case of a misspecified model is studied in section 4. Finally, some conclusions are presented in section 5.

2 Framework

The aim is to estimate the total $t_y = \sum_U y_k$ of one study variable $\boldsymbol{y}' = (y_1, y_2, \cdots, y_N)$ in a population U with unit labels $\{1, 2, \cdots, N\}$ where N is known. It is assumed that there is one auxiliary variable $\boldsymbol{x}' = (x_1, x_2, \cdots, x_N)$, $x_k > 0$, known for each element in U. A without-replacement sample s of size n is selected and y_k is observed for all units $k \in s$.

In this section we shall describe the six strategies that are spanned by two designs, stratified simple random sampling —STSI— and proportional-to-size sampling —πps— on the one hand; and three estimators, the Horvitz-Thompson estimator —HT—, the poststratified estimator —pos— and the regression estimator —reg— on the other hand.

The reasoning behind these strategies is as follows. Regarding the design, simple random sampling does not make any use of the auxiliary information, whereas πps makes, what we call, strong use of it. STSI lies in between, we will say that it makes weak use of the auxiliary information. In a similar way, regarding the estimator, the HT estimator does not make use of the auxiliary information, as opposed to the reg-estimator that makes strong use of it. The pos-estimator lies in between, making weak use of the auxiliary information. Then the six strategies make use of the auxiliary information at a different degree.

The general regression estimator —GREG— is described in the first part of this section. The HT, pos and reg estimators are shown to be particular cases of it. In the last part of the section, the superpopulation model is described.

Before moving on, a note on notation is convenient. Throughout the paper we will use the symbols E and E for expectation and model residuals, respectively.

2.1 The GREG estimator

In the general setting, we have J auxiliary variables, i.e. the vector $\boldsymbol{x}_k = (x_{1k}, x_{2k}, \cdots, x_{Jk})$ is available for each $k \in U$. The GREG estimator of t_y is defined as

$$\hat{t}_{\text{GREG}} \equiv \sum_U \hat{y}_k + \sum_s \frac{e_{ks}}{\pi_k},$$

where π_k is the inclusion probability of the kth element, $e_{ks} = y_k - \hat{y}_k$ and $\hat{y}_k = \boldsymbol{x}_k \widehat{\boldsymbol{B}}$ with

$$\widehat{\boldsymbol{B}} = \left(\sum_s \frac{\boldsymbol{x}_k' \boldsymbol{x}_k}{a_k \pi_k} \right)^{-1} \sum_s \frac{\boldsymbol{x}_k' y_k}{a_k \pi_k} \tag{1}$$

The a-values will be defined later. No closed expression for the variance of the GREG estimator is available, but it can be approximated by (see Särndal et al., 1992)

$$\text{AV}_{\text{p}}\left[\hat{t}_{\text{GREG}}\right] = \sum_U \sum_U (\pi_{kl} - \pi_k \pi_l) \frac{e_k}{\pi_k} \frac{e_l}{\pi_l} \qquad \text{with } e_k = y_k - \boldsymbol{x}_k \boldsymbol{B} \tag{2}$$

where π_{kl} is the second order inclusion probability of k and l and

$$\boldsymbol{B} = \left(\sum_U \frac{\boldsymbol{x}_k' \boldsymbol{x}_k}{a_k} \right)^{-1} \sum_U \frac{\boldsymbol{x}_k' y_k}{a_k}.$$

This is the same expression for the variance of the HT estimator with e_k instead of y_k. From now on we will write $\text{V}_{\text{p}}\left[\hat{t}_{\text{GREG}}\right]$ instead of $\text{AV}_{\text{p}}\left[\hat{t}_{\text{GREG}}\right]$, i.e. we assume that the approximation exactly coincides with the variance.

The following are sufficient but not necessary conditions for (2) being equal to zero:

i. $e_k = 0$ for all $k \in U$. The e_k depend only on the estimator, not the design, therefore a GREG estimator that correctly explains the study variable will lead to small residuals. (In this case, not only the approximation but the true variance is equal to zero, and the GREG estimator is exactly equal to t_y.)
ii. $\pi_k = n\, e_k/t_e$ with $t_e = \sum_U e_k$. Even if the e_k were known, this condition cannot be fulfilled as some residuals will be smaller than zero while some will be larger than zero, thus leading to negative probabilities.
iii. $\pi_k = n\frac{|e_k|}{t_{|e|}}$ together with $\pi_{kl} = \pi_k\pi_l$ if $k \in U^+$ and $l \in U^-$, where $t_{|e|} = \sum_U |e_k|$ $U^+ = \{k, e_k \geq 0\}$ and $U^- = \{k, e_k < 0\}$. One method to satisfy the second part of the condition would be to stratify the population with respect to the sign of e_k which, however, requires knowledge about the finite population at a level of detail that is seldom available. We will therefore assume that this knowledge is not available and we will settle for the next condition.
iii'. $\pi_k = n\frac{|e_k|}{t_{|e|}}$, which is obtained if we drop the $\pi_{kl} = \pi_k\pi_l$ part of condition **iii**. Note that **iii'** does not yield a zero variance. Why consider condition **iii'** then? First, as will be shown below, the HT estimator can be seen as a particular case of the GREG estimator and if we have $y_k > 0$, it is equivalent to condition **ii** above, thus leading to a zero variance. Second, it will be useful for defining the so-called optimal strategy and model-based stratification.

As can be seen, in the context of the GREG estimator, conditions **i** and **iii'** suggest the specific role of the design and the estimator in the sampling strategy. The estimator must explain the trend of the study variable with respect to the auxiliary variable, leading to small residuals. The design, on the other hand, must explain the residuals, in other words, how the study variable is spread around the trend.

The Horvitz-Thompson estimator as a particular case of the GREG estimator Consider the case where the auxiliary vector is of the form $\boldsymbol{x}_k = 0$ for all $k \in U$. If we allow $0/0 = 0$ (this terrible blasphemy is justified by using a generalized inverse in (1) instead of the inverse, and noting that 0 is a generalized inverse of itself) we have that

$$\widehat{\boldsymbol{B}} = \left(\sum_s \frac{\boldsymbol{x}_k'\boldsymbol{x}_k}{a_k\pi_k}\right)^{-} \sum_s \frac{\boldsymbol{x}_k' y_k}{a_k\pi_k} = 0$$

then $\hat{y}_k = \boldsymbol{x}_k\widehat{\boldsymbol{B}} = 0$ and $e_{ks} = y_k - \hat{y}_k = y_k - 0 = y_k$. The GREG estimator becomes

$$\hat{t}_{\text{GREG}} = \sum_U \hat{y}_k + \sum_s \frac{e_{ks}}{\pi_k} = \sum_U 0 + \sum_s \frac{y_k}{\pi_k} = \hat{t}_\pi$$

which explicitly shows that the HT estimator can be seen as the case where no auxiliary information is used into the GREG estimator. Note also that $e_k = y_k - \boldsymbol{x}_k\boldsymbol{B} = y_k$, therefore (2) becomes the exact variance of the HT estimator.

The poststratified estimator Let $U_1', U_2' \cdots, U_G'$ be a partition of U. Consider the case where the auxiliary vector is of the form $\boldsymbol{x}_k = (x_{1k}, x_{2k}, \cdots, x_{Gk})$ with x_{gk} defined as

$$x_{gk} = \begin{cases} 1 & \text{if } k \in U_g' \\ 0 & \text{otherwise} \end{cases}$$

This means that the auxiliary information for each element is a vector that indicates a group (poststratum) to which the element belongs.

The poststratified estimator, or simply pos-estimator, is obtained when this particular type of auxiliary information is used in the GREG estimator. The residuals become

$$e_k = y_k - B_g \qquad \text{with} \qquad B_g = \frac{t_{y/a,g}}{t_{1/a,g}} \qquad (k \in U'_g)$$

where $t_{y/a,g} = \sum_{U'_g} \frac{y_k}{a_k}$ and $t_{1/a,g} = \sum_{U'_g} \frac{1}{a_k}$. We will consider the case where a_k is constant within poststrata, $a_k = c_g$, then $B_g = \bar{y}_{U'_g}$.

The regression estimator Consider the case where the auxiliary vector is of the form $\boldsymbol{x}_k = (1, z_k)$, with z_k the result of a known function applied to the known x_k. The regression estimator, or simply reg-estimator, is obtained when this $\boldsymbol{x}_k$ is used in the GREG estimator. The residuals become

$$e_k = y_k - (B_0 + B_1 z_k) \text{ with } B_1 = \frac{t_{1/a} t_{zy/a} - t_{z/a} t_{y/a}}{t_{1/a} t_{z^2/a} - t^2_{z/a}} \text{ and } B_0 = \frac{t_{y/a}}{t_{1/a}} - B_1 \frac{t_{z/a}}{t_{1/a}}$$

where $t_{1/a} = \sum_U \frac{1}{a_k}$, $t_{y/a} = \sum_U \frac{y_k}{a_k}$, $t_{z/a} = \sum_U \frac{z_k}{a_k}$, $t_{z^2/a} = \sum_U \frac{z_k^2}{a_k}$ and $t_{zy/a} = \sum_U \frac{z_k y_k}{a_k}$. We will consider the case where $a_k = c$, then $B_1 = \frac{N t_{zy} - t_z t_y}{N t_{z^2} - t_z^2}$ and $B_0 = \frac{t_y}{N} - B_1 \frac{t_z}{N}$, where $t_y = \sum_U y_k$, $t_z = \sum_U z_k$, $t_{z^2} = \sum_U z_k^2$ and $t_{zy} = \sum_U z_k y_k$.

2.1.1 The GREG estimator and STSI

In STSI the population U is partitioned (stratified) into H groups (strata) denoted U_h, $h = 1, \cdots, H$, with sizes N_h. In each stratum a simple random sample, s_h, of a predefined size n_h is selected. Under STSI sampling, the (approximation to the) variance of the GREG estimator becomes

$$\mathrm{V}_{\mathrm{STSI}}\left[\hat{t}_{\mathrm{GREG}}\right] = \sum_{h=1}^{H} \frac{N_h^2}{n_h}\left(1 - \frac{n_h}{N_h}\right) S^2_{eU_h} \tag{3}$$

where $S^2_{eU_h} = \frac{1}{N_h - 1} \sum_{U_h} (e_k - \bar{e}_{U_h})^2$, with e_k as defined above and $\bar{e}_{U_h} = \frac{1}{N_h} \sum_{U_h} e_k$.

According to Dalenius and Hodges (1959) there are four operations that must be defined when using stratified sampling: **i.** the choice of the stratification variable; **ii.** the choice of the number of strata, H; **iii.** the boundaries of the strata; and, **iv.** the allocation of the sample size, n, into the strata. For the purposes of this paper, the first operation is not under discussion: all we have is $\boldsymbol{x}$. We will also let H to be arbitrarily defined. For the third operation, we will use the approximation to the cum$\sqrt{f}$-rule as described by Särndal et al. (1992). Finally, Neyman optimal allocation will be used for the fourth operation.

2.1.2 The GREG estimator and πps

A sampling design satisfying the following conditions will be called a *strict* πps: **i.** being a without-replacement design; **ii.** having a fixed sample size ($\sum_U \pi_k = n$); **iii.** the inclusion probabilities induced by the design, π_k, coincide with some desired inclusion probabilities, π_k^*; **iv.** second order inclusion probabilities strictly larger than zero, $\pi_{kl} > 0 \ \forall k, l \in U$; **v.** π_{kl} easy to compute; **vi.** selection scheme easy to implement for any sample size $n = 1, \cdots, N$.

In the literature we find many designs that satisfy some but not all the conditions above. Hanif and Brewer (1980) and Tillé (2006), for example, present reviews of available designs. Rosén (1997) introduces Pareto πps, which satisfies the conditions above except **iii.** and **v**. However, the difference between the actual inclusion probabilities and the desired ones is negligible (Rosén, 2000b). Also, approximate expressions for π_{kl} are available. Therefore, Pareto πps will be the πps considered in this paper.

Under πps, the (approximation to the) variance of the GREG estimator becomes (Rosén, 2000a)

$$\mathrm{V}_{\pi\mathrm{ps}}\left[\hat{t}_{\mathrm{GREG}}\right] = \frac{N}{N-1}\left[t_{e^2(1-\pi^*)/\pi^*} - \frac{t^2_{e(1-\pi^*)}}{t_{\pi^*(1-\pi^*)}}\right]$$

where $t_{e^2(1-\pi^*)/\pi^*} = \sum_U e_k^2(1-\pi_k^*)/\pi_k^*$, $t_{e(1-\pi^*)} = \sum_U e_k(1-\pi_k^*)$ and $t_{\pi^*(1-\pi^*)} = \sum_U \pi_k^*(1-\pi_k^*)$, with e_k as defined above.

2.2 The superpopulation model and the strategies under comparison

At the beginning of this section six sampling strategies were mentioned. Five of them will be defined here in the frame of a superpopulation model. The reasons for not considering the remaining one will be given.

We will assume that when defining the sampling strategy, the statistician is willing to admit that the following model *adequately describes* the relation between the study variable, $\boldsymbol{y}$, and the auxiliary variable, $\boldsymbol{x}$. The values of the study variable $\boldsymbol{y}$ are realizations of the model ξ_0

$$Y_k = \delta_0 + \delta_1 x_k^{\delta_2} + \epsilon_k \tag{4}$$

The error terms ϵ_k are random variables satisfying

$$\mathrm{E}_{\xi_0}\left[\epsilon_k\right] = 0 \qquad \mathrm{V}_{\xi_0}\left[\epsilon_k\right] = \delta_3^2 x_k^{2\delta_4} \qquad \mathrm{E}_{\xi_0}\left[\epsilon_k \epsilon_l\right] = 0 \ (k \neq l)$$

where the moments are taken with respect to the model ξ_0, and δ_i are constant parameters.

It is worth recalling that this model is considered at the planning stage of the survey, when no y-values are available. Therefore it is not possible to consider the estimation of the δ-parameters and the best that can be done is to propose some guess or to consider some values taken from previous studies.

The term $\delta_0 + \delta_1 x_k^{\delta_2}$ in model ξ_0 will be called *trend*, where δ_0 is the intercept, δ_2 is the shape and δ_1 is a scale factor. The term $\delta_3^2 x_k^{2\delta_4}$ will be called *spread*, where δ_4 is the shape and δ_3 is a scale factor. Brewer (1963; 2002, p. 111 and p. 200-201) shows rather heuristically that for most survey data $1/2 \leq \delta_4 \leq 1$ when $\delta_2 = 1$.

Model ξ_0 as defined above is then used for assisting the definition of the sampling strategy as follows.

Strategy 1, πps(δ_4)–reg(δ_2) At the design stage consider πps with $\pi_k = n\frac{x_k^{\delta_4}}{t_{x^{\delta_4}}}$. At the estimation stage consider the reg-estimator with $\boldsymbol{x}_k = (1, x_k^{\delta_2})$.

Justification If model ξ_0 is assumed, it is natural to consider the GREG estimator with $\boldsymbol{x}_k = (1, x_k^{\delta_2})$ at the estimation stage. In this case, we have

$$y_k = B_0 + B_1 x_k^{\delta_2} + e_k \qquad \text{but also} \qquad y_k = \delta_0 + \delta_1 x_k^{\delta_2} + \epsilon_k^*$$

where e_k is the residual resulting from fitting the regression underlying the GREG estimator and ϵ_k^* is a realization of the random variable ϵ_k. Then, for large populations (so that convergence for B_0 and B_1 has been approximately achieved), we have

$$e_k = (\delta_0 - B_0) + (\delta_1 - B_1) x_k^{\delta_2} + \epsilon_k^* \approx \epsilon_k^*$$

In order to minimize the variance in the sense of condition **iii'** one would like to use a design having $\pi_k = n \frac{|e_k|}{t_{|e|}}$. Using the approximation above, we get

$$|e_k| \approx |\epsilon_k^*| = \sqrt{\epsilon_k^{*2}} \approx \sqrt{\mathrm{E}_{\xi_0}\left[\epsilon_k^2\right]} = \sqrt{\delta_3^2 x_k^{2\delta_4}} = \delta_3 x_k^{\delta_4}$$

Therefore the design must satisfy $\pi_k = n \frac{x_k^{\delta_4}}{t_{x^{\delta_4}}}$.

A comprehensive definition of this strategy can be found in, for example, Särndal et al. (1992). This strategy is often found in the literature and referred to as "optimal", in the sense that it minimizes an approximation to the anticipated variance, $\mathrm{E}_{\xi_0}\mathrm{V}_{\mathrm{p}}[\hat{t}]$, a model dependent statistic.

Strategy 2, STSI(δ_4)–reg(δ_2) At the design stage consider STSI with strata defined by using the cum$\sqrt{f}$-rule on $x_k^{\delta_4}$ and Neyman allocation. At the estimation stage consider the reg-estimator with $\boldsymbol{x}_k = (1, x_k^{\delta_2})$.

Justification Assuming the model ξ_0, the GREG estimator with $\boldsymbol{x}_k = (1, x_k^{\delta_2})$ is used again and we get $|e_k| \approx \delta_3 x_k^{\delta_4}$. Ignoring the factor δ_3, the strata are then constructed using the approximation to the cum$\sqrt{f}$-rule on $x_k^{\delta_4}$ together with Neyman allocation.

This strategy, known as model-based stratification, was proposed by Wright (1983), who also showed a lower bound for its efficiency compared to πps(δ_4)–reg(δ_2). For a comprehensive description, see, for example, Särndal et al. (1992, section 12.4).

Strategy 3, STSI(δ_2)–HT At the design stage consider STSI with strata defined by using the cum$\sqrt{f}$-rule on $x_k^{\delta_2}$ and Neyman allocation. At the estimation stage consider the HT estimator.

Justification As mentioned above, the HT estimator can be seen as the case when null auxiliary information is used in the GREG estimator. In this case the residuals are $e_k = y_k$ and in order to have a small variance (3) we look for strata leading to a small sum-of-squares-within, $SSW_y = \sum_{h=1}^{H} \sum_{U_h} (y_k - \bar{y}_{U_h})^2$.

Using the model, a proxy for y_k is $y_k \approx \delta_0 + \delta_1 x_k^{\delta_2}$, which leads to

$$SSW_y = \sum_{h=1}^{H} \sum_{U_h} (y_k - \bar{y}_{U_h})^2 \approx \delta_1^2 \sum_{h=1}^{H} \sum_{U_h} \left(x_k^{\delta_2} - \overline{x_{U_h}^{\delta_2}} \right)^2 = \delta_1^2 \, SSW_{x^{\delta_2}} \tag{5}$$

So we have to look for strata leading to a small SSW of $x_k^{\delta_2}$. The strata are then created using the approximation to the cum$\sqrt{f}$-rule on $x_k^{\delta_2}$ together with Neyman allocation.

The first two strategies make use of the auxiliary information at both the design and the estimation stage. On the other hand, the strategy that couples STSI with the HT estimator uses auxiliary information only at the design stage in a way that we call weak. This strategy will be considered as a benchmark.

Strategy 4, πps(δ_4)–pos(δ_2) At the design stage consider πps with $\pi_k = n\frac{x_k^{\delta_4}}{t_{x^{\delta_4}}}$. At the estimation stage consider the pos-estimator with poststrata defined by using the cum$\sqrt{f}$-rule on $x_k^{\delta_2}$.

Justification It is worth justifying the reason for considering this strategy. On one hand, the regression estimator makes an explicit assumption of an underlying model ξ_0, which in practice will almost certainly not be fully correct. On the other hand, the HT estimator completely ignores the available auxiliary information. The poststratified estimator can be seen as a compromise between those two scenarios.

In this case we have two decisions to make, namely, how will the poststrata be defined in order to have small residuals, e_k, and how will the inclusion probabilities be defined in order to explain the resulting residuals. Regarding the first task, recall that the residuals of the pos-estimator can be written as $e_k = y_k - \bar{y}_{U'_g}$ for all $k \in U'_g$, where $\bar{y}_{U'_g}$ is the average of the y-values in the gth poststratum. When looking for poststrata that minimize these e_k, a natural criterion would be to minimize its square sum, $\sum_U e_k^2$, but note that

$$\sum_U e_k^2 = \sum_{g=1}^{G} \sum_{U'_g} e_k^2 = \sum_{g=1}^{G} \sum_{U'_g} \left(y_k - \bar{y}_{U'_g}\right)^2$$

which is the SSW shown in (5) above. Therefore we use the same approach, and the poststrata will be created using the approximation to the cum$\sqrt{f}$-rule on $x_k^{\delta_2}$.

Regarding the second task, we use an approach analogous to the one considered for πps–reg. Note that $y_k = B_g + e_k$ but also $y_k = \delta_0 + \delta_1 x_k^{\delta_2} + \epsilon_k^*$ where e_k is the residual resulting from fitting the poststratification estimator and ϵ_k^* is a realization of the random variable ϵ_k. Then

$$e_k = \delta_0 + \delta_1 x_k^{\delta_2} + \epsilon_k^* - B_g$$

In order to minimize the variance in the sense of condition **iii'** one would like to use a design having $\pi_k = n\frac{|e_k|}{t_{|e|}}$. As the e_k are unknown, we use the following approximation

$$|e_k| = |\delta_0 + \delta_1 x_k^{\delta_2} + \epsilon_k^* - B_g| = \sqrt{\left(\delta_0 + \delta_1 x_k^{\delta_2} - B_g + \epsilon_k^*\right)^2} \approx$$
$$\sqrt{\mathrm{E}_{\xi_0}\left[\left(\delta_0 + \delta_1 x_k^{\delta_2} - B_g + \epsilon_k\right)^2\right]} \approx \sqrt{\left(\delta_0 + \delta_1 x_k^{\delta_2} - B_g\right)^2 + \mathrm{E}_{\xi_0}[\epsilon_k^2]} \approx \delta_3 x_k^{\delta_4}$$

The first approximation uses the expected value of the random variable ϵ_k as an approximation to a realization from it; the second approximation assumes that convergence has been achieved for B_g; and $x_k^{\delta 2} \approx \overline{x_{U'_g}^{\delta_2}}$ was used in order to obtain the last expression. Using condition **iii'** and these proxies for the residuals, we have that the design must satisfy $\pi_k = n\frac{x_k^{\delta_4}}{t_{x^{\delta_4}}}$.

Strategy 5, STSI(δ_4)–pos(δ_2) At the design stage consider STSI with strata defined by using the cum$\sqrt{f}$-rule on $x_k^{\delta_4}$ and Neyman allocation. At the estimation stage consider the pos-estimator with poststrata defined by using the cum$\sqrt{f}$-rule on $x_k^{\delta_2}$.

Justification In this case the poststratified estimator is used again in the same way as in the strategy above, which means that poststrata are created using the approximation to the cum$\sqrt{f}$-rule on $x_k^{\delta_2}$. The same approximated residuals are then obtained.

The strata are defined by applying the approximation to the cum$\sqrt{f}$-rule on x^{δ_4} and the sample is allocated using Neyman allocation.

A simulation study by Rosén (2000a) suggests that, for $\delta_2 = 1$ and $1/2 \leq \delta_4 < 1$, πps sampling with the GREG estimator is better than πps sampling with the HT estimator. This is an argument for not considering the strategy πps–HT any longer.

3 Simulation study under a correctly specified model

In this section we will assume that the model considered by the statistician holds, i.e. the y-values are realizations of the model ξ_0

$$Y_k = \delta_0 + \delta_1 x_k^{\delta_2} + \epsilon_k \qquad \text{with } \mathrm{E}_{\xi_0}[\epsilon_k] = 0 \qquad \mathrm{V}_{\xi_0}[\epsilon_k] = \delta_3^2 x_k^{2\delta_4} \qquad \mathrm{E}_{\xi_0}[\epsilon_k \epsilon_l] = 0 \ (k \neq l) \tag{6}$$

We will compare the performance of the five strategies under different conditions. As mentioned in the last section, πps(δ_4)–reg(δ_2) is expected to perform the best.

Under the model (6), the design variance becomes a random variable as it varies with every finite population generated by the superpopulation model. Therefore, we will say that the most efficient strategy is the one that yields the smallest expectation $\mathrm{E}_{\xi_0}\mathrm{V}_{\mathrm{p}}[\hat{t}]$, the anticipated variance. Closed expressions for this value are not easily obtained, therefore we appeal to a simulation study, defined as follows.

1. The auxiliary variable $\boldsymbol{x}$ is generated as N realizations from a gamma distribution with shape equal to $\alpha = \frac{4}{\gamma^2}$ and scale $\lambda = 12\gamma^2$, where γ is the desired skewness, plus one unit. In this way we have $\mathrm{E}[X] = \frac{4}{\gamma^2} \cdot 12\gamma^2 + 1 = 49$.

2. y_k are realizations from $Y_k = \delta_0 + \delta_1 x_k^{\delta_2} + \epsilon_k$ with $\epsilon_k \sim N(0, \delta_3^2 x_k^{2\delta_4})$.

3. The design variance of a sample of size n is then computed for each strategy.

4. Steps 1 to 3 are repeated $R = 5000$ times.

5. The anticipated variance for each strategy is approximated as the mean of the R replicates of the design variance, i.e. $\mathrm{E}_{\xi_0}\mathrm{V}_{\mathrm{p}}\left[\hat{t}\right] \approx \frac{1}{R}\sum_{r=1}^{R} \mathrm{V}_{\mathrm{p}}^{(r)}[\hat{t}] \equiv \bar{\mathrm{V}}_{\mathrm{p}}[\hat{t}]$.

The simulation depends on several factors (the size of the finite population, N; the skewness of X, γ; the sample size, n; the parameters in the model, δ_i). In addition, the number of strata and poststrata, H and G, must be specified for four strategies. The following values (levels) were considered:

- The population size was fixed at $N = 5000$ and the sample size at $n = 500$, thus obtaining a fixed sampling fraction of $f = n/N = 0.1$.
- Two levels of skewness were considered: moderate ($\gamma = 3$) and high ($\gamma = 12$).
- The number of strata/poststrata was fixed at $H = G = 5$.
- Only the case with no intercept, $\delta_0 = 0$, will be studied. Three values for the trend shape are considered: $\delta_2 = 0.75$, 1 and 1.25 (concave, linear and convex association, respectively). Also three values for the spread shape are considered: $\delta_4 = 0.5$, 0.75 and 1 (low, moderate and high heteroscedasticity, respectively).
- As mentioned by Rosén (2000a), one of the two parameters δ_1 or δ_3 is redundant. Therefore we consider only the case $\delta_1 = 1$. The value of δ_3 required for obtaining a given Pearson's Correlation Coefficient —PCC—, ρ, is

$$\delta_3^2 = \frac{\lambda^{2(\delta_2-\delta_4)}}{\Gamma(\alpha+2\delta_4)}\left[\frac{\left(\Gamma(\alpha+1+\delta_2)-\alpha\Gamma(\alpha+\delta_2)\right)^2}{\alpha\Gamma(\alpha)\rho^2}-\Gamma(\alpha+2\delta_2)+\frac{\Gamma^2(\alpha+\delta_2)}{\Gamma(\alpha)}\right], \tag{7}$$

  where $\Gamma(\cdot)$ is the gamma function and α and λ as defined above. Given all the other parameters, we found the values of δ_3 required for obtaining a desired PCC of $\rho = 0.65$ and 0.95 (moderate and high correlation respectively).

The simulation defined in this way leads to $36 = 2 \times 3 \times 3 \times 2$ (two levels for γ, three levels for δ_2, three levels for δ_4 and two levels for ρ) scenarios. Table 1 shows the simulated expected variance $\mathrm{E}_{\xi_0}\mathrm{V}_\mathrm{p}[\hat{t}]$ of each strategy in each scenario. The results are shown as a percentage of the expected variance of STSI(δ_2)–HT, which is shown in the column "Reference". The rows are sorted from the scenario that yields the least gain with respect to STSI(δ_2)–HT to the one yielding the largest gain. Bold values indicate the most efficient strategy in each scenario. The main results are summarized as follows:

- As expected, the strategies using auxiliary information at both stages are in general more efficient than the reference.
- No strategy was always more efficient than STSI(δ_2)–HT. However, STSI(δ_4)–reg(δ_2) and πps(δ_4)–pos(δ_2) were better in almost every scenario. In fact they yield the best results in most scenarios where $\gamma = 12$ and $\delta_4 \geq 0.75$.
- πps(δ_4)–reg(δ_2) was the most efficient strategy in most scenarios. This is, however, not a surprise as it is supposed to be optimal. What comes as a surprise is the fact that it is not always the best. This is explained by the fact that it minimizes an approximation to the anticipated variance, not the anticipated variance itself. Its optimality relies on several assumptions, like the model being correct (which is true in this case) and the population size being so large that B_0 and B_1 have essentially no variance. When the simulations are run with $N = 300000$ (results not shown), πps(δ_4)–reg(δ_2) becomes indeed the best in every scenario.
- It is worth to remark that although asymptotically optimal, πps(δ_4)–reg(δ_2) might be quite inefficient in highly skewed or highly heteroscedastic populations even when the model is correct.

Table 1: Simulated $E_{\xi_0} V_p[\hat{t}]$ as a percentage of the anticipated variance of STSI(δ_2)–HT

γ	ρ	δ_2	δ_4	Reference	πps–reg	STSI–reg	πps–pos	STSI-pos
3	0.65	0.75	0.50	$1.32 \cdot 10^7$	**84.7**	98.7	88.9	102.7
3	0.65	1.00	0.50	$2.43 \cdot 10^8$	**80.3**	93.7	83.6	97.3
3	0.65	1.00	0.75	$1.63 \cdot 10^8$	**77.7**	97.6	82.9	101.7
3	0.65	1.25	0.75	$2.76 \cdot 10^9$	**76.7**	96.4	80.2	100.5
3	0.65	0.75	0.75	$9.61 \cdot 10^6$	**75.2**	94.4	83.2	100
3	0.65	1.25	1.00	$1.79 \cdot 10^9$	**74.8**	100.2	81.0	104.8
3	0.65	1.25	0.50	$4.51 \cdot 10^9$	**72.6**	84.7	75.4	88.4
3	0.65	1.00	1.00	$1.14 \cdot 10^8$	**70.1**	93.7	82.3	100
12	0.65	0.75	0.50	$1.14 \cdot 10^7$	**68.7**	83.5	69.6	84.2
3	0.65	0.75	1.00	$7.32 \cdot 10^6$	**62.3**	83.3	82.6	91.5
12	0.95	1.25	1.00	$1.43 \cdot 10^7$	218.3	81.6	**62.2**	350.7
3	0.95	1.00	0.50	$2.58 \cdot 10^7$	**59.8**	69.7	91.4	104.5
3	0.95	1.25	0.50	$3.64 \cdot 10^8$	**56.3**	65.6	91.5	112.8
12	0.95	0.75	0.50	$6.64 \cdot 10^5$	**53.9**	65.5	74.0	76.7
3	0.95	1.25	0.75	$2.55 \cdot 10^8$	**52.1**	65.4	92.5	110.6
3	0.95	1.00	0.75	$1.95 \cdot 10^7$	**51.5**	64.7	97.7	100
12	0.65	0.75	0.75	$1.06 \cdot 10^6$	51.7	86.0	**51.3**	100
3	0.95	0.75	0.50	$1.28 \cdot 10^6$	**50.8**	59.3	96.1	101.1
12	0.65	1.00	0.75	$5.40 \cdot 10^7$	51.2	88.3	**47.6**	93.4
3	0.95	1.25	1.00	$1.94 \cdot 10^8$	**43.1**	57.8	115.9	100.9
12	0.95	1.00	0.75	$5.10 \cdot 10^6$	**43.0**	73.6	59.2	128.0
3	0.95	1.00	1.00	$1.56 \cdot 10^7$	**40.5**	54.2	142.2	100
12	0.65	1.00	1.00	$4.84 \cdot 10^6$	261.8	82.2	**39.8**	100
3	0.95	0.75	0.75	$1.07 \cdot 10^6$	**39.4**	49.4	114.1	100
12	0.95	1.25	0.75	$2.03 \cdot 10^8$	**39.2**	68.8	44.8	106.3
12	0.65	1.25	1.00	$1.86 \cdot 10^8$	295.9	113.0	**38.4**	134.0
12	0.65	1.25	0.75	$3.57 \cdot 10^9$	40.0	70.5	**37.0**	72.6
12	0.65	1.00	0.50	$1.13 \cdot 10^9$	**36.0**	43.8	36.2	44.0
12	0.95	1.00	0.50	$9.02 \cdot 10^7$	**35.8**	43.5	39.5	46.2
3	0.95	0.75	1.00	$9.37 \cdot 10^5$	**28.4**	37.9	199.4	102.6
12	0.65	0.75	1.00	$2.89 \cdot 10^5$	96.8	28.0	**33.1**	79.4
12	0.95	1.00	1.00	$1.21 \cdot 10^6$	83.2	**26.0**	64.3	100
12	0.95	1.25	0.50	$4.87 \cdot 10^9$	**24.5**	29.8	26.7	31.9
12	0.65	1.25	0.50	$8.88 \cdot 10^{10}$	**24.4**	29.7	24.4	29.8
12	0.95	0.75	0.75	$1.93 \cdot 10^5$	**13.0**	21.8	49.8	100
12	0.95	0.75	1.00	$1.57 \cdot 10^5$	8.3	**2.3**	46.2	98.6

4 The case of a misspecified model

In the previous section we verified empirically that when the finite population is generated by the model ξ_0, πps(δ_4)–reg(δ_2) is in fact the best among the strategies being compared. In this section we will study how robust the results are when the model is misspecified. In the first part we will define the type of misspecification that will be studied in the paper. The results of a simulation study will be presented in section 4.2. In section 4.3, expressions for approximating the anticipated variance will be presented. These expressions are assessed in section 4.4.

4.1 The misspecified model

First, we will define how "misspecification"shall be understood in this paper. ξ_0 (which from now on will be called *working model*) reflects the knowledge or beliefs the statistician has about the relation between $\boldsymbol{x}$ and $\boldsymbol{y}$ at the design stage. Nevertheless, one hardly believes that this is the true generating model. We will assume that this true model exists but it is unknown to the statistician. It will be denoted by ξ. Any deviation of ξ_0 with respect to ξ is a misspecification of the model. As this definition is too wide and in order to keep the analysis tractable, we will limit ourselves to a very simple type of misspecification, which is when the working model is of the form (4) or (6) and the true model, ξ, is

$$Y_k = \beta_0 + \beta_1 x_k^{\beta_2} + \epsilon_k \qquad \text{with } \mathrm{E}_\xi[\epsilon_k] = 0 \qquad \mathrm{V}_\xi[\epsilon_k] = \beta_3^2 x_k^{2\beta_4} \qquad \mathrm{E}_\xi[\epsilon_k \epsilon_l] = 0 \ (k \neq l)$$

with $\beta_2 \neq \delta_2$ or $\beta_4 \neq \delta_4$.

4.2 Simulation study under the misspecified model

A simulation study was carried out in order to compare the performance of the five strategies under this type of misspecification. The results are divided into three groups. The first one, when the trend term is correct ($\delta_2 = \beta_2$) but the spread is misspecified ($\delta_4 \neq \beta_4$). The second one, when the spread term is correct ($\delta_4 = \beta_4$) but the trend is misspecified ($\delta_2 \neq \beta_2$). The last case is when both, trend and spread, are misspecified ($\delta_2 \neq \beta_2$ and $\delta_4 \neq \beta_4$).

The setup is similar to the one used in the simulations in section 3. The only difference being that now y_k are realizations from $Y_k = \beta_0 + \beta_1 x_k^{\beta_2} + \epsilon_k$ with $\epsilon_k \sim N(0, \beta_3^2 x_k^{2\beta_4})$. Now, the most efficient strategy is the one that yields the smallest anticipated variance under ξ, $\mathrm{E}_\xi \mathrm{V}_\mathrm{p}[\hat{t}]$.

Regarding the factors, we set $N = 5000$, $n = 500$, $H = 5$, $\beta_0 = 0$, $\beta_1 = 1$, $\gamma = 3, 12$, $\beta_2 = 0.75, 1, 1.25$ and $\beta_4 = 0.5, 0.75, 1$. β_3 as defined in (7) replacing δ_2 and δ_4 by β_2 and β_4, respectively. The strategies are defined using $\delta_2 = 0.75, 1, 1.25$ and $\delta_4 = 0.5, 0.75, 1$.

Table 2 shows the results for the 72 scenarios in the case of correct trend but misspecified spread. The results are shown as a percentage of the expected variance of STSI(δ_2)–HT. The scenarios are sorted from the one that yields the least gain with respect to STSI(δ_2)–HT to the one yielding the largest gain. Bold values indicate the most efficient strategy in each scenario. The absence of a bold value indicates that STSI(δ_2)–HT was the most efficient strategy. The main results are summarized as follows:

- There were several cases where STSI(δ_2)–HT was the most efficient strategy.
- Although πps(δ_4)–reg(δ_2) was still the best strategy in most scenarios, there were many cases were it was overcome by either STSI(δ_4)–reg(δ_2) or πps(δ_4)–pos(δ_2). Unlike the simulation in section 2, results do not get better when the population size is increased.

Table 2: Simulated $E_\xi V_p[\hat{t}]$ in the case of correct trend and misspecified spread.

γ	ρ	δ_2	β_4	δ_4	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.65	1.25	1.00	0.50	356.0	357.5	361.3	412.7
12	0.65	1.00	1.00	0.50	275.2	257.4	286.6	305.3
12	0.95	1.25	1.00	0.50	254.3	257.5	1019.8	997.4
12	0.65	0.75	0.50	1.00	166.9	189.9	166.7	191.2
12	0.95	0.75	0.50	1.00	130.0	149.6	140.0	172.2
12	0.95	1.25	1.00	0.75	115.9	146.2	164.4	682.8
3	0.65	1.25	1.00	0.50	105.4	137.4	112.1	147.1
3	0.65	0.75	0.50	1.00	140.8	102.3	152.6	106.9
3	0.65	1.00	1.00	0.50	**98.6**	128.8	105.4	136.6
3	0.65	1.00	0.50	1.00	133.6	**97.1**	139.9	100
3	0.65	0.75	0.50	0.75	96.5	**95.9**	102.4	100
12	0.65	0.75	0.50	0.75	**94.8**	98.7	95.0	100
3	0.65	1.00	0.50	0.75	91.5	**91.0**	95.1	93.8
3	0.65	1.25	0.50	1.00	120.7	**87.8**	123.6	89.6
3	0.65	0.75	1.00	0.50	**87.6**	114.3	95.1	121.5
12	0.95	1.00	0.50	1.00	**87.3**	99.0	87.7	100
3	0.65	1.00	0.75	1.00	**86.7**	95.6	95.7	100
12	0.65	1.00	0.50	1.00	86.7	99.9	**86.4**	100
3	0.65	1.00	0.75	0.50	**86.4**	110.2	91.1	115.6
12	0.65	1.00	0.75	0.50	88.9	**85.9**	91.9	90.3
3	0.65	1.25	0.75	1.00	**85.6**	94.5	90.0	97.4
3	0.65	1.25	0.75	0.50	**85.3**	108.8	89.7	115.0
12	0.65	0.75	1.00	0.50	99.4	**84.9**	117.7	110.1
3	0.65	0.75	0.75	1.00	**83.9**	92.5	99.8	98.7
12	0.65	0.75	0.75	0.50	88.2	**83.5**	95.6	90.3
3	0.65	0.75	0.75	0.50	**83.5**	106.5	89.3	112.0
3	0.65	1.25	0.50	0.75	82.8	**82.3**	84.9	84.8
3	0.65	1.25	1.00	0.75	**80.6**	111.3	85.6	117.7
12	0.95	1.00	1.00	0.50	85.8	**80.4**	350.7	275.8
3	0.65	1.00	1.00	0.75	**75.5**	104.2	82.6	110.2
12	0.95	0.75	0.50	0.75	**74.3**	77.6	85.1	100
3	0.95	1.00	0.50	1.00	99.5	**72.3**	161.4	100
12	0.95	1.00	0.75	0.50	74.8	**72.1**	140.9	119.9
12	0.65	1.25	0.75	0.50	69.9	**68.3**	71.2	71.1
3	0.95	1.25	0.50	1.00	93.6	**68.1**	132.6	91.0
3	0.95	1.00	0.50	0.75	68.3	**67.9**	103.1	94.5
3	0.65	0.75	1.00	0.75	**67.0**	92.6	77.3	100
12	0.95	1.25	0.75	0.50	68.6	**67.0**	123.5	118.6
12	0.65	0.75	0.75	1.00	72.9	96.1	**64.7**	110.1
3	0.95	1.25	0.50	0.75	64.1	**63.7**	92.4	95.3
12	0.65	1.25	1.00	0.75	162.7	203.0	**63.6**	244.5
3	0.95	0.75	0.50	1.00	84.5	**61.5**	210.2	108.8
12	0.95	1.00	0.75	1.00	**61.1**	82.6	63.5	100

Continued on next page

Table 2 – *Continued from previous page*

γ	ρ	δ_2	β_4	δ_4	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.65	1.00	0.75	1.00	72.5	98.4	**60.9**	100
3	0.95	1.25	1.00	0.50	**60.8**	79.4	126.7	167.7
12	0.65	1.00	1.00	0.75	143.5	147.7	**60.8**	204.1
12	0.95	1.25	0.50	1.00	59.3	68.2	**59.1**	68.9
12	0.65	1.25	0.50	1.00	59.1	68.0	**58.9**	68.0
3	0.95	1.25	0.75	1.00	**58.1**	64.1	113.6	96.8
3	0.95	1.25	0.75	0.50	**58.0**	73.9	108.0	140.9
3	0.95	0.75	0.50	0.75	57.9	**57.6**	120.7	100
3	0.95	1.00	0.75	1.00	**57.4**	63.3	139.5	100
3	0.95	1.00	0.75	0.50	**57.2**	72.9	98.9	118.8
3	0.95	1.00	1.00	0.50	**57.0**	74.3	109.1	131.8
12	0.65	1.00	0.50	0.75	50.0	51.9	**49.9**	52.1
12	0.95	1.00	0.50	0.75	**49.7**	51.6	50.7	54.7
12	0.95	1.25	0.75	1.00	55.5	77.1	**49.4**	96.1
12	0.65	1.25	0.75	1.00	56.5	78.7	**48.0**	79.7
3	0.95	1.25	1.00	0.75	**46.6**	64.3	99.6	123.9
12	0.95	1.00	1.00	0.75	**45.3**	46.1	99.9	272.1
3	0.95	0.75	0.75	1.00	**43.9**	48.4	194.0	105.0
3	0.95	0.75	0.75	0.50	**43.7**	55.7	97.7	105.6
3	0.95	1.00	1.00	0.75	**43.6**	60.2	101.1	104.5
3	0.95	0.75	1.00	0.50	**39.9**	52.1	101.6	109.1
12	0.65	0.75	1.00	0.75	56.4	48.8	**38.6**	100
12	0.95	1.25	0.50	0.75	**33.7**	35.1	33.9	36.6
12	0.65	1.25	0.50	0.75	33.7	35.0	**33.6**	35.1
3	0.95	0.75	1.00	0.75	**30.6**	42.2	116.0	100
12	0.95	0.75	0.75	0.50	22.4	**21.2**	92.1	60.4
12	0.95	0.75	0.75	1.00	**18.3**	24.0	51.4	102.8
12	0.95	0.75	1.00	0.50	8.2	**7.0**	92.4	54.4
12	0.95	0.75	1.00	0.75	4.6	**4.1**	48.4	100

Values are shown as a percentage of the expected variance of STSI(δ_2)–HT. Bold values indicate the most efficient strategy in each scenario. The absence of a bold value indicates that STSI(δ_2)–HT was the most efficient strategy.

Table 3 shows the results for the 72 scenarios in the case of correct spread but misspecified trend. It can be seen that, except in a few scenarios, πps(δ_4)–reg(δ_2) is no longer the best strategy. In fact, it becomes the worst one in most scenarios, sometimes with a variance more than ten times bigger than that of any other strategy.

Table 3: Simulated $\mathrm{E}_\xi \mathrm{V}_\mathrm{p}[\hat{t}]$ in the case of correct spread and misspecified trend.

γ	ρ	δ_4	δ_2	β_2	πps–reg	STSI–reg	πps–pos	STSI-pos
3	0.65	1.00	1.25	0.75	244.7	**91.7**	127.6	98.8

Continued on next page

Table 3 – *Continued from previous page*

γ	ρ	δ_4	δ_2	β_2	πps–reg	STSI–reg	πps–pos	STSI-pos
3	0.65	1.00	1.25	1.00	106.0	97.3	**91.6**	102.8
3	0.65	1.00	1.00	0.75	121.0	**90.9**	109.9	100
3	0.65	0.50	0.75	1.25	95.7	100.5	**88.9**	104.6
3	0.65	0.50	0.75	1.00	88.9	99.8	**88.7**	103.7
3	0.65	0.75	1.00	0.75	89.6	95.0	**88.6**	100
3	0.65	0.75	1.25	0.75	116.2	90.1	**87.9**	94.3
3	0.65	0.50	1.00	1.25	**83.3**	94.6	84.1	98.8
3	0.65	0.50	1.00	0.75	**82.9**	92.0	83.8	95.3
3	0.65	0.75	1.25	1.00	83.5	94.4	**81.6**	98.4
3	0.65	0.75	1.00	1.25	86.0	98.7	**81.6**	102.7
3	0.65	0.75	0.75	1.00	87.0	95.8	**79.8**	100
3	0.95	0.50	0.75	1.25	164.0	**79.0**	96.4	120.5
3	0.65	0.75	0.75	1.25	107.2	96.4	**78.6**	100
3	0.65	1.00	1.00	1.25	99.6	95.5	**74.9**	100
3	0.65	0.50	1.25	0.75	81.8	81.3	**74.8**	83.5
3	0.95	0.50	0.75	1.00	93.4	**74.4**	92.8	109.5
3	0.65	0.50	1.25	1.00	**74.1**	83.8	74.8	86.6
3	0.95	0.50	1.00	1.25	88.6	**73.9**	99.9	124.5
12	0.95	0.50	0.75	1.00	102.0	83.5	**72.5**	85.1
12	0.95	0.50	0.75	1.25	188.9	88.4	**72.1**	86.2
12	0.65	0.50	0.75	1.00	72.3	84.7	**69.6**	84.8
12	0.65	0.50	0.75	1.25	76.2	84.8	**69.3**	84.6
3	0.65	1.00	0.75	1.00	101.6	85.5	**68.9**	90.6
3	0.95	0.75	1.00	1.25	131.3	**68.2**	88.0	110.7
3	0.95	0.75	0.75	1.25	323.8	**68.1**	73.3	100
3	0.95	0.75	0.75	1.00	138.4	**65.8**	82.9	100
3	0.65	1.00	0.75	1.25	172.0	86.8	**65.0**	90.0
3	0.95	0.50	1.25	1.00	74.5	**63.0**	81.5	89.0
3	0.95	1.00	0.75	1.25	879.6	65.6	**61.2**	90.3
3	0.95	0.75	1.25	1.00	113.5	**60.8**	100	91.7
3	0.95	1.00	1.00	1.25	297.8	**59.5**	91.1	100
12	0.95	1.00	1.00	1.25	5555.1	153.9	**58.7**	100
3	0.95	1.00	0.75	1.00	318.9	**57.7**	90.2	95.1
12	0.95	1.00	1.25	1.00	2461.9	64.7	**54.3**	216.3
3	0.95	0.50	1.00	0.75	83.8	**53.9**	94.0	85.5
3	0.95	1.00	1.25	1.00	258.7	**53.8**	173.1	90.3
12	0.95	0.75	1.00	1.25	367.5	98.0	**53.8**	121.6
12	0.95	0.75	0.75	1.00	456.1	126.4	**52.9**	100
12	0.95	0.75	0.75	1.25	1741.9	241.9	**52.1**	100
12	0.65	0.75	0.75	1.00	101.1	103.0	**51.9**	100
12	0.65	0.75	0.75	1.25	164.0	109.1	**51.8**	100
12	0.95	0.75	1.00	0.75	292.5	**50.3**	53.3	108.7
12	0.95	0.75	1.25	1.00	222.6	68.1	**48.4**	115.4
12	0.65	0.75	1.00	0.75	94.5	83.0	**47.6**	92.9
3	0.95	0.75	1.00	0.75	145.7	**47.4**	140.3	86.7

Continued on next page

Table 3 – *Continued from previous page*

γ	ρ	δ_4	δ_2	β_2	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.65	0.75	1.00	1.25	70.8	90.4	**46.9**	91.7
12	0.95	1.00	1.00	0.75	687.2	48.6	**46.6**	100
3	0.95	0.50	1.25	0.75	136.4	**46.5**	80.5	63.9
3	0.95	1.00	1.00	0.75	369.9	**42.5**	307.7	100
12	0.65	1.00	1.00	0.75	583.9	64.7	**42.3**	100
3	0.95	1.00	1.25	0.75	862.0	**42.2**	314.7	75.7
3	0.95	0.75	1.25	0.75	306.0	**41.8**	133.9	68.4
12	0.65	1.00	1.25	1.00	829.9	103.3	**40.7**	141.3
12	0.95	0.75	1.25	0.75	684.6	62.2	**39.3**	94.6
12	0.95	0.50	1.00	0.75	68.3	**39.1**	45.9	47.9
12	0.95	0.50	1.00	1.25	47.5	44.3	**38.4**	45.7
12	0.65	0.75	1.25	1.00	56.7	70.5	**37.6**	74.8
12	0.65	0.75	1.25	0.75	145.1	69.3	**36.9**	74.6
12	0.65	0.50	1.00	0.75	37.9	43.7	**36.5**	44.2
12	0.65	0.50	1.00	1.25	36.7	43.8	**36.0**	43.9
12	0.65	1.00	1.00	1.25	801.8	105.8	**34.0**	100
12	0.95	1.00	0.75	1.00	2133.4	191.9	**32.2**	61.6
12	0.65	1.00	1.25	0.75	1377.9	81.0	**31.8**	98.6
12	0.95	1.00	1.25	0.75	1756.6	70.0	**30.5**	93.5
12	0.95	0.50	1.25	0.75	93.8	**28.5**	32.4	34.7
12	0.95	0.50	1.25	1.00	30.5	29.8	**27.5**	32.2
12	0.65	0.50	1.25	0.75	28.0	29.7	**24.8**	30.0
12	0.65	0.50	1.25	1.00	24.9	29.7	**24.6**	29.9
12	0.95	1.00	0.75	1.25	13444.2	745.2	**23.3**	51.1
12	0.65	1.00	0.75	1.00	650.6	89.7	**21.0**	57.5
12	0.65	1.00	0.75	1.25	1494.4	126.0	**18.2**	55.0

Values are shown as a percentage of the expected variance of STSI(δ_2)–HT. Bold values indicate the most efficient strategy in each scenario.

Table 4 shows the results for the 144 scenarios in the case of misspecified trend and spread. Again, πps(δ_4)–reg(δ_2) is no longer the best strategy; in fact, it is the worst one, with variances more than ten times bigger than that of any other strategy in several scenarios. On the other hand, STSI(δ_4)–reg(δ_2) and πps(δ_4)–pos(δ_2) are now the best options.

Table 4: Simulated $\mathrm{E}_\xi \mathrm{V}_\mathrm{p}[\hat{t}]$ in the case of misspecified trend and spread.

γ	ρ	δ_2	β_2	δ_4	β_4	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.65	1.25	1.00	0.50	1.00	486.8	301.9	349.4	378.0
12	0.65	1.00	1.25	0.50	1.00	510.6	315.7	296.0	337.9
12	0.95	1.00	1.25	0.50	1.00	2153.3	247.9	580.8	518.0
12	0.95	0.75	1.25	0.50	1.00	5489.4	290.0	230.6	192.0

Continued on next page

Table 4 – *Continued from previous page*

γ	ρ	δ_2	β_2	δ_4	β_4	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.65	0.75	1.00	1.00	0.50	175.4	191.8	167.8	191.3
12	0.95	0.75	1.25	1.00	0.50	461.3	205.8	166.9	190.2
12	0.65	0.75	1.25	1.00	0.50	183.8	193.3	166.3	192.4
12	0.95	0.75	1.00	1.00	0.50	246.9	193.9	164.4	188.7
12	0.65	0.75	1.25	0.50	1.00	716.5	187.6	155.9	176.8
12	0.65	0.75	1.00	0.50	1.00	374.5	157.7	153.8	165.7
12	0.95	1.25	1.00	0.50	1.00	880.2	117.1	497.0	434.5
12	0.65	1.00	0.75	0.50	1.00	294.4	109.3	158.9	151.4
12	0.95	1.00	1.25	0.75	1.00	2778.8	212.0	108.7	405.3
12	0.95	0.75	1.25	0.50	0.75	1402.4	132.6	121.8	108.2
3	0.65	1.25	1.00	0.50	1.00	108.0	133.5	109.4	140.6
12	0.65	1.25	0.75	0.50	1.00	565.8	108.0	150.3	148.6
3	0.65	1.25	0.75	0.50	1.00	122.9	123.5	106.3	128.7
3	0.65	1.00	1.25	0.50	1.00	105.1	131.1	106.8	140.3
3	0.65	0.75	1.25	1.00	0.50	201.9	104.2	142.8	105.9
3	0.65	0.75	1.00	1.00	0.50	163.2	103.6	145.4	106.4
3	0.65	1.00	0.75	0.50	1.00	103.0	123.5	104.9	130.5
12	0.65	0.75	1.25	0.50	0.75	185.0	100.1	98.6	**98.4**
3	0.65	1.00	1.25	1.00	0.50	147.5	**97.9**	136.6	100
3	0.65	0.75	1.25	0.75	0.50	119.7	**97.4**	99.1	100
3	0.65	0.75	1.00	0.75	0.50	105.1	**96.9**	99.9	100
3	0.65	1.25	0.75	0.75	1.00	138.2	100.3	**96.5**	106.6
12	0.65	0.75	1.00	0.50	0.75	133.9	**96.4**	99.5	97.8
12	0.95	0.75	1.25	0.75	0.50	249.4	113.0	**95.8**	100
3	0.65	1.00	0.75	1.00	0.50	156.6	**95.7**	151.6	100
12	0.65	0.75	1.00	0.75	0.50	99.5	100.2	**95.6**	100
12	0.65	0.75	1.25	0.75	0.50	104.4	100.7	**95.4**	100
3	0.65	0.75	1.25	0.50	1.00	107.6	117.8	**95.3**	125.2
3	0.65	0.75	1.00	0.50	1.00	95.2	116.5	**94.8**	123.5
12	0.95	0.75	1.00	0.75	0.50	137.4	102.8	**94.7**	100
3	0.65	1.00	0.75	1.00	0.75	122.6	**93.6**	114.8	100
3	0.65	1.25	1.00	1.00	0.75	105.8	**92.8**	96.8	96.4
3	0.65	1.00	1.25	0.75	0.50	97.2	**91.8**	94.3	94.5
12	0.95	1.25	1.00	0.75	1.00	1130.6	**91.4**	110.5	367.8
3	0.65	1.00	1.25	0.50	0.75	**90.7**	111.6	91.9	118.0
3	0.65	1.00	1.25	1.00	0.75	107.2	96.9	**90.6**	100
3	0.65	1.00	0.75	0.75	1.00	92.3	100.3	**90.6**	107.4
12	0.95	1.00	1.25	0.50	0.75	316.6	**89.9**	132.8	120.8
3	0.65	1.00	0.75	0.50	0.75	**89.8**	107.2	91.3	112.1
3	0.65	0.75	1.00	1.00	0.75	114.2	94.2	**89.5**	98.1
3	0.65	1.00	0.75	0.75	0.50	99.2	**89.4**	98.6	92.8
3	0.65	1.25	0.75	1.00	0.75	200.2	**89.4**	120.4	94.2
3	0.65	0.75	1.25	0.50	0.75	98.7	109.0	**89.3**	114.6
3	0.65	0.75	1.00	0.50	0.75	89.2	108.0	**88.9**	113.3
12	0.95	1.00	1.25	1.00	0.50	120.8	100.3	**88.0**	100

Continued on next page

Table 4 – *Continued from previous page*

γ	ρ	δ_2	β_2	δ_4	β_4	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.65	1.00	1.25	0.50	0.75	104.1	**87.8**	89.8	89.5
3	0.65	1.25	1.00	0.75	1.00	91.1	108.2	**87.7**	114.2
12	0.95	0.75	1.00	0.50	1.00	930.4	**87.7**	177.9	129.5
3	0.65	1.25	0.75	0.50	0.75	98.5	101.5	**87.5**	104.9
3	0.65	1.25	1.00	0.50	0.75	**87.2**	106.6	88.2	111.1
12	0.65	1.00	0.75	1.00	0.50	92.2	99.6	**87.0**	100
12	0.65	1.00	1.25	1.00	0.50	89.1	100	**86.9**	100
3	0.65	1.25	1.00	1.00	0.50	132.8	**86.8**	127.3	89.0
3	0.65	0.75	1.25	1.00	0.75	167.0	95.1	**86.4**	97.5
12	0.95	0.75	1.00	0.50	0.75	400.2	**85.2**	121.8	101.1
3	0.65	1.25	0.75	1.00	0.50	191.0	**84.8**	141.0	87.9
3	0.95	0.75	1.25	1.00	0.50	700.0	**82.5**	111.5	100.4
12	0.95	1.00	0.75	1.00	0.50	172.7	92.7	**81.9**	100
3	0.65	1.25	1.00	0.75	0.50	86.9	**81.4**	85.8	83.8
3	0.95	1.00	1.25	0.50	1.00	104.7	**81.3**	123.2	165.2
3	0.65	1.00	1.25	0.75	1.00	87.4	106.0	**80.7**	111.8
3	0.95	0.75	1.00	1.00	0.50	293.6	**79.3**	137.7	104.9
3	0.95	0.75	1.25	0.50	0.75	179.7	**79.3**	98.2	129.4
3	0.95	0.75	1.25	0.50	1.00	196.6	**79.1**	102.4	136.7
3	0.65	1.25	0.75	0.75	0.50	107.4	**79.0**	89.5	81.7
3	0.95	1.00	1.25	0.50	0.75	95.3	**78.3**	110.2	145.6
12	0.65	1.00	0.75	0.50	0.75	116.1	**77.4**	90.2	86.1
3	0.95	1.00	1.25	1.00	0.50	255.1	**75.4**	130.5	100
3	0.95	0.75	1.00	0.50	0.75	94.9	**73.7**	94.0	117.2
3	0.95	0.75	1.25	0.75	0.50	295.0	**73.5**	87.2	100
3	0.95	0.75	1.25	1.00	0.75	784.8	**73.0**	73.6	94.5
3	0.65	0.75	1.00	0.75	1.00	82.5	94.4	**72.8**	100
3	0.95	1.25	1.00	0.50	1.00	92.0	**72.7**	104.3	118.6
3	0.95	0.75	1.00	0.75	0.50	141.5	**72.4**	96.8	100
3	0.95	0.75	1.00	0.50	1.00	99.4	**72.3**	98.3	123.8
3	0.65	0.75	1.25	0.75	1.00	108.9	95.3	**71.3**	100
3	0.95	1.00	1.25	0.75	0.50	128.4	**70.5**	95.8	102.5
3	0.95	1.25	1.00	0.50	0.75	82.7	**69.1**	92.5	105.3
12	0.65	1.25	1.00	0.50	0.75	81.6	**68.4**	74.1	73.2
12	0.65	1.25	1.00	0.75	1.00	398.5	178.0	**68.2**	246.5
3	0.95	0.75	1.00	1.00	0.75	292.6	**67.7**	100.1	99.3
12	0.65	0.75	1.25	1.00	0.75	291.0	121.7	**67.5**	110.6
3	0.95	1.00	1.25	1.00	0.75	263.1	**67.5**	98.2	100
12	0.65	0.75	1.00	1.00	0.75	161.6	115.5	**66.7**	110.2
3	0.95	1.25	1.00	1.00	0.50	211.7	**64.8**	164.2	85.5
3	0.95	1.00	1.25	0.75	1.00	143.2	**64.8**	89.1	117.8
12	0.95	0.75	1.25	1.00	0.75	3406.4	275.2	**63.7**	104.1
12	0.95	1.25	1.00	0.50	0.75	193.7	**63.4**	129.2	116.4
3	0.95	0.75	1.25	0.75	1.00	357.3	**63.2**	69.0	100
12	0.95	1.00	1.25	1.00	0.75	724.9	106.4	**62.6**	100

Continued on next page

Table 4 – *Continued from previous page*

γ	ρ	δ_2	β_2	δ_4	β_4	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.65	1.25	0.75	0.50	0.75	142.1	**62.3**	71.9	69.4
12	0.65	1.00	1.25	1.00	0.75	111.2	100.4	**60.7**	100
3	0.95	1.25	1.00	0.75	0.50	107.7	**60.6**	98.1	82.9
3	0.95	1.25	1.00	1.00	0.75	224.9	**60.2**	159.5	88.8
12	0.95	0.75	1.00	1.00	0.75	835.2	147.9	**59.9**	99.1
3	0.95	0.75	1.00	0.75	1.00	145.2	**59.5**	79.5	100
12	0.65	1.25	1.00	1.00	0.50	60.6	67.4	**58.9**	67.5
3	0.95	1.00	0.75	1.00	0.50	317.9	**58.9**	274.2	100
12	0.65	1.00	0.75	1.00	0.75	156.7	93.0	**58.7**	100
12	0.95	1.25	1.00	1.00	0.50	76.9	67.1	**58.7**	68.3
12	0.65	1.25	0.75	1.00	0.50	68.8	67.4	**58.5**	67.5
3	0.95	1.25	1.00	0.75	1.00	125.1	**58.4**	108.3	97.7
12	0.95	1.25	0.75	1.00	0.50	254.0	67.8	**55.8**	70.4
12	0.65	1.00	0.75	0.75	1.00	304.7	84.8	**52.9**	136.4
3	0.95	1.00	0.75	0.75	0.50	141.2	**52.8**	137.0	85.2
3	0.95	1.00	0.75	0.50	0.75	87.0	**52.5**	99.5	91.1
3	0.95	1.25	0.75	1.00	0.50	670.2	**52.2**	269.9	76.8
12	0.65	1.00	1.25	0.75	1.00	399.8	182.1	**51.7**	198.3
3	0.95	1.00	0.75	0.50	1.00	91.6	**50.4**	105.9	94.8
12	0.95	1.00	0.75	0.75	0.50	93.5	**50.3**	50.9	60.7
3	0.95	1.00	0.75	1.00	0.75	337.3	**50.1**	283.8	100
12	0.65	1.00	0.75	0.75	0.50	52.5	51.9	**50.0**	52.4
12	0.95	1.00	0.75	1.00	0.75	558.3	59.5	**50.0**	100
12	0.95	1.00	1.25	0.75	0.50	65.1	52.3	**49.9**	53.4
12	0.65	1.00	1.25	0.75	0.50	50.5	51.8	**49.5**	51.8
12	0.95	1.25	1.00	1.00	0.75	446.5	74.2	**49.0**	99.9
12	0.95	0.75	1.00	0.75	1.00	1156.7	170.9	**48.6**	100
12	0.65	1.25	1.00	1.00	0.75	91.4	78.3	**47.9**	80.6
3	0.95	1.25	0.75	1.00	0.75	773.9	**47.1**	290.1	76.8
3	0.95	1.25	0.75	0.50	0.75	153.8	**46.5**	86.2	67.6
12	0.65	1.25	0.75	1.00	0.75	276.2	77.6	**45.6**	80.5
3	0.95	1.25	0.75	0.50	1.00	167.6	**45.2**	91.1	69.2
3	0.95	1.25	0.75	0.75	0.50	269.6	**44.8**	127.3	66.8
12	0.65	1.25	0.75	0.75	1.00	697.7	98.0	**44.0**	129.7
3	0.95	1.00	0.75	0.75	1.00	155.3	**42.2**	149.1	87.4
12	0.95	1.00	0.75	0.75	1.00	362.6	**41.9**	55.4	116.6
12	0.95	0.75	1.25	0.75	1.00	6874.0	675.7	**40.3**	100
12	0.95	1.25	0.75	0.75	1.00	853.2	61.0	**39.6**	102.3
3	0.95	1.25	0.75	0.75	1.00	337.4	**38.2**	142.9	68.2
12	0.95	1.25	0.75	1.00	0.75	1380.7	71.8	**34.7**	89.7
12	0.95	1.25	0.75	0.75	0.50	129.9	37.8	**34.4**	42.7
12	0.95	1.25	1.00	0.75	0.50	42.4	35.3	**34.3**	37.4
12	0.65	1.25	1.00	0.75	0.50	34.5	35.2	**33.8**	35.3
12	0.65	1.25	0.75	0.75	0.50	38.5	35.2	**33.5**	35.4
12	0.65	0.75	1.00	0.75	1.00	350.5	118.8	**31.2**	100

Continued on next page

Table 4 – *Continued from previous page*

γ	ρ	δ_2	β_2	δ_4	β_4	πps–reg	STSI–reg	πps–pos	STSI-pos
12	0.95	1.25	0.75	0.50	0.75	498.9	**30.8**	88.6	72.5
12	0.95	1.00	0.75	0.50	0.75	232.5	**27.6**	104.6	77.2
12	0.65	0.75	1.25	0.75	1.00	762.8	158.5	**27.3**	100
12	0.95	1.25	0.75	0.50	1.00	636.7	**24.7**	100.2	79.8
12	0.95	1.00	0.75	0.50	1.00	275.7	**16.5**	113.0	79.7

Values are shown as a percentage of the expected variance of STSI(δ_2)–HT. Bold values indicate the most efficient strategy in each scenario. The absence of a bold value indicates that STSI(δ_2)–HT was the most efficient strategy.

The simulations in this section yield evidence that the so-called optimal strategy, πps(δ_4)–reg(δ_2), is not robust towards misspecifications of the model. This fact will be reinforced analytically by means of the approximations to the anticipated variance that are shown in the following section.

4.3 The anticipated variance

In this section, approximated expressions for the anticipated variance, $\mathrm{E}_\xi \mathrm{V}_\mathrm{p}[\hat{t}]$, will be obtained for each strategy defined in section 2.2. The following notation will be used: $\overline{x^c_A} = \frac{1}{N_A}\sum_A x^c_k$, $S_{c,d,A} = \frac{1}{N_A}\sum_A (x^c_k - \overline{x^c_A})(x^d_k - \overline{x^d_A})$, where c and d are constants, A is a subset of U and N_A its cardinality. If A is not indicated, then $A = U$.

Before starting with the approximations, the following result is required. In it, an approximation of β_3 in terms of β_1 and the correlation between $\boldsymbol{x}$ and $\boldsymbol{y}$, $R_{x,y}$, is established. The proof of this result can be found in the Appendix.

Result 1. Under ξ,

$$\beta_3^2 \approx \beta_1^2 F_0 \qquad \text{with } F_0 = \frac{S^2_{1,\beta_2} - R^2_{x,y} S_{1,1} S_{\beta_2,\beta_2}}{R^2_{x,y}\overline{x^{2\beta_4}} S_{1,1}} \tag{8}$$

where $R_{x,y} = \frac{\sum_U (x_k - \overline{x}_U)(y_k - \overline{y}_U)}{\sqrt{\sum_U (x_k - \overline{x}_U)^2 \sum_U (y_k - \overline{y}_U)^2}}$ is the finite population coefficient of correlation between $\boldsymbol{x}$ and $\boldsymbol{y}$.

The following result establishes an approximation to the anticipated variance for each of the five strategies defined in section 2.2.

Result 2. If ξ_0 is assumed when ξ is the true model, then:

1. the anticipated variance of πps(δ_4)–reg(δ_2) can be approximated by

$$\mathrm{AA}_{\xi,\pi\mathrm{ps}}\left[\hat{t}_{\mathrm{reg}}\right] \equiv \beta_1^2 \frac{N^2 \overline{x^{\delta_4}}}{n}\left(\left(S_{\beta_2,\beta_2-\delta_4} - \overline{x^{\beta_2}} S_{\beta_2,-\delta_4}\right) - 2\frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}}(S_{\beta_2,\delta_2-\delta_4} - \overline{x^{\delta_2}} S_{\beta_2,-\delta_4}) + \frac{S^2_{\beta_2,\delta_2}}{S^2_{\delta_2,\delta_2}}(S_{\delta_2,\delta_2-\delta_4} - \overline{x^{\delta_2}} S_{\delta_2,-\delta_4}) + F_0 \overline{x^{2\beta_4-\delta_4}}\right) \tag{9}$$

2. the anticipated variance of STSI(δ_4)–reg(δ_2) can be approximated by

$$\mathrm{AA}_{\xi,\mathrm{STSI}}\left[\hat{t}_{\mathrm{reg}}\right] \equiv \beta_1^2 \sum_h \frac{N_h^2}{n_h}\left(\left(S_{\beta_2,\beta_2,U_h} - 2\frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} S_{\beta_2,\delta_2,U_h} + \frac{S_{\beta_2,\delta_2}^2}{S_{\delta_2,\delta_2}^2} S_{\delta_2,\delta_2,U_h}\right) + F_0 \overline{x_{U_h}^{2\beta_4}}\right) \quad (10)$$

Note that if we let $v_k = x_k^{\beta_2} - x_k^{\delta_2}\frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}}$, (10) can be written as

$$\mathrm{AA}_{\xi,\mathrm{STSI}}\left[\hat{t}_{\mathrm{reg}}\right] \equiv \beta_1^2 \sum_h \frac{N_h^2}{n_h}\left(S_{v,U_h}^2 + F_0 \overline{x_{U_h}^{2\beta_4}}\right) \qquad \text{with} \qquad S_{v,U_h}^2 = \frac{1}{N_h}\sum_{U_h}\left(v_k - \overline{v}_{U_h}\right)^2$$

3. the anticipated variance of STSI(δ_2)–HT can be approximated by

$$\mathrm{AA}_{\xi,\mathrm{STSI}}\left[\hat{t}_{\mathrm{HT}}\right] \equiv \beta_1^2 \sum_h \frac{N_h^2}{n_h}\left(S_{\beta_2,\beta_2,U_h} + F_0 \overline{x_{U_h}^{2\beta_4}}\right) \quad (11)$$

4. the anticipated variance of πps(δ_4)–pos(δ_2) can be approximated by

$$\mathrm{AA}_{\xi,\pi\mathrm{ps}}\left[\hat{t}_{\mathrm{pos}}\right] \equiv \beta_1^2 \frac{N\overline{x^{\delta_4}}}{n}\left(\sum_g N_g\left(S_{\beta_2,\beta_2-\delta_4,U_g} - \overline{x_{U_g}^{\beta_2}} S_{\beta_2,-\delta_4,U_g}\right) + N F_0 \overline{x^{2\beta_4-\delta_4}}\right) \quad (12)$$

5. Let $U_{hg} = U_h \cap U'_g$ be the intersection between the hth stratum and the gth poststratum. The anticipated variance of STSI(δ_4)–pos(δ_2) can be approximated by

$$\mathrm{AA}_{\xi,\mathrm{STSI}}\left[\hat{t}_{\mathrm{pos}}\right] \equiv \beta_1^2 \sum_h \frac{N_h^2}{n_h}\left(\frac{1}{N_h}\sum_g N_{hg}\left(\overline{x_{U_{hg}}^{2\beta_2}} - 2\overline{x_{U'_g}^{\beta_2}}\,\overline{x_{U_{hg}}^{\beta_2}} + \overline{x_{U'_g}^{\beta_2}}^2\right) - \frac{1}{N_h^2}\left(\sum_g N_{hg}\left(\overline{x_{U_{hg}}^{\beta_2}} - \overline{x_{U'_g}^{\beta_2}}\right)\right)^2 + F_0 \overline{x_{U_h}^{2\beta_4}}\right) \quad (13)$$

It can be seen that even under this simple misspecification of the model, πps(δ_4)–reg(δ_2) does not minimize the approximation to the anticipated variance. Let us compare, for example, (9) and (10) in the case where $\delta_2 = \beta_2$. In that case we get

$$\mathrm{AA}_{\xi,\pi\mathrm{ps}}\left[\hat{t}_{\mathrm{reg}}\right] = \beta_1^2 \frac{N^2}{n} F_0 \overline{x^{\delta_4}}\,\overline{x^{2\beta_4-\delta_4}} \qquad \text{and} \qquad \mathrm{AA}_{\xi,\mathrm{STSI}}\left[\hat{t}_{\mathrm{reg}}\right] = \beta_1^2 F_0 \sum_h \frac{N_h^2}{n_h}\overline{x_h^{2\beta_4}}$$

If we allow $S_{\delta_4,\delta_4,h}$ constant, and take into account that Neyman optimal allocation was used, we get that $\mathrm{AA}_{\xi,\pi\mathrm{ps}}\left[\hat{t}_{\mathrm{reg}}\right] > \mathrm{AA}_{\xi,\mathrm{STSI}}\left[\hat{t}_{\mathrm{reg}}\right]$ when $2\beta_4 < \delta_4$ and $\delta_4 > 0$.

We will not focus on this type of comparisons, instead we will empirically assess how good these approximations are.

4.4 Assessing the approximations to the anticipated variance

The expressions in result 2 are asymptotic as they assume $N \to \infty$, they also assume n/N is small. It is then natural to ask how good are these approximations for finite values of N. Therefore, in order to assess them, the simulations were run once more with $N = 50000$, $n = 50$ and $R = 200$. Every time $\boldsymbol{x}$ is generated, the approximation to the anticipated variance is computed and compared with the anticipated variance obtained by simulation, as follows

$$D^{(r)} = \frac{\left|\mathrm{AA}_{\xi,\mathrm{p}}^{(r)}\left[\hat{t}\right] - \mathrm{E}_{\xi}^{*}\mathrm{V}_{\mathrm{p}}^{(r)}\left[\hat{t}\right]\right|}{\mathrm{E}_{\xi}^{*}\mathrm{V}_{\mathrm{p}}^{(r)}\left[\hat{t}\right]}.$$

Table 5 shows the percentage of cases where $D^{(r)}$ was less than 0.05 for each strategy in each scenario. The rows have been sorted from the one that yields the worst results to the best one. It can be seen that the expressions in result 2 approximate the intended anticipated variance very well in most cases. Some care must be taken when the skewness is large, however.

Table 5: Percentage of cases where $D^{(r)}$ was less than 0.05.

γ	ρ	β_2	δ_2	β_4	δ_4	πps–reg	STSI–reg	STSI–HT	πps–pos	STSI-pos
12	0.65	1.25	1.25	0.50	1.00	41.0	40.5	41.5	41.0	41.5
12	0.65	1.25	0.75	0.50	0.75	47.0	41.0	42.0	42.0	42.0
12	0.65	1.25	0.75	0.50	0.50	45.5	42.0	42.0	44.5	43.5
12	0.65	1.25	1.00	0.50	1.00	45.0	44.5	44.5	45.0	44.5
12	0.65	1.25	1.00	0.50	0.50	45.5	45.5	45.0	46.5	46.0
12	0.65	1.25	1.25	0.50	0.75	46.0	46.0	45.5	46.5	47.5
12	0.65	1.25	0.75	0.50	1.00	51.0	46.0	46.5	45.5	46.0
12	0.65	1.25	1.00	0.50	0.75	47.5	46.5	47.0	47.0	47.5
12	0.65	1.25	1.25	0.50	0.50	47.5	47.5	48.0	48.5	48.5
12	0.95	1.25	1.25	0.50	0.75	53.0	53.0	57.0	60.0	63.5
12	0.95	1.25	1.25	0.50	1.00	52.0	52.5	60.5	56.0	66.0
12	0.65	1.25	1.25	0.75	0.50	59.0	58.0	59.5	61.5	61.0
12	0.65	1.25	0.75	0.75	0.50	65.5	56.0	58.0	59.0	61.0
12	0.65	1.25	1.25	0.75	0.75	60.0	59.0	62.0	60.0	60.0
12	0.65	1.25	1.00	0.75	1.00	77.0	57.0	60.0	59.0	60.0
12	0.65	1.25	0.75	0.75	1.00	95.5	54.5	55.0	53.5	55.5
12	0.65	1.25	1.25	0.75	1.00	62.0	62.0	64.5	63.5	65.0
12	0.65	1.00	1.25	0.50	0.75	64.0	63.0	63.0	63.5	64.5
12	0.65	1.00	1.25	0.50	1.00	64.5	64.5	63.0	62.5	64.5
12	0.65	1.25	1.00	0.75	0.50	63.5	62.0	65.0	65.0	67.0
12	0.95	1.25	1.00	0.50	1.00	76.5	58.0	66.0	57.5	66.0
12	0.65	1.00	1.00	0.50	1.00	65.0	66.0	66.5	65.5	66.5
12	0.65	1.25	0.75	0.75	0.75	80.5	62.0	63.5	62.0	63.5
12	0.65	1.00	0.75	0.50	0.50	67.5	65.5	65.0	68.0	66.0
12	0.65	1.00	1.00	0.50	0.50	65.5	66.5	66.5	69.5	67.0
12	0.95	1.25	1.25	0.50	0.50	57.0	55.5	64.0	81.0	77.5
12	0.65	1.25	1.00	0.75	0.75	71.5	65.0	68.5	64.0	66.5
12	0.65	1.00	1.00	0.50	0.75	66.5	68.0	68.0	68.0	68.5
12	0.65	1.00	0.75	0.50	1.00	68.0	68.5	70.0	66.0	69.0
12	0.95	1.25	1.00	0.50	0.75	77.5	56.5	68.0	67.0	73.5
12	0.95	0.75	0.75	1.00	0.75	38.5	33.0	88.5	94.0	90.0
12	0.95	1.00	1.00	0.50	1.00	62.0	64.5	76.5	65.5	76.0
12	0.95	0.75	0.75	1.00	0.50	40.0	39.0	92.0	88.5	86.0
12	0.65	1.00	0.75	0.50	0.75	69.5	68.0	69.5	69.0	69.5
12	0.95	1.00	1.00	0.50	0.75	64.0	64.5	72.0	71.0	75.5
12	0.65	1.00	1.25	0.50	0.50	69.5	68.5	70.5	70.0	69.5
12	0.95	1.25	1.00	0.50	0.50	69.0	56.0	65.0	80.5	77.5
12	0.65	1.25	1.25	1.00	1.00	62.0	74.5	78.0	65.5	79.0
12	0.95	1.25	1.25	1.00	1.00	68.0	66.5	90.5	49.5	89.5
12	0.95	0.75	0.75	0.75	0.75	43.0	39.5	94.5	95.0	94.5
12	0.95	0.75	0.75	1.00	1.00	41.0	39.5	87.5	100	99.5
12	0.95	1.25	1.25	0.75	1.00	68.0	67.0	79.5	74.5	80.5

Continued on next page

Table 5 – *Continued from previous page*

γ	ρ	β_2	δ_2	β_4	δ_4	πps–reg	STSI–reg	STSI–HT	πps–pos	STSI-pos
12	0.95	1.25	1.25	0.75	0.75	65.5	63.5	80.5	74.0	87.5
12	0.95	1.25	1.00	1.00	1.00	100	80.5	84.5	26.0	88.5
12	0.95	0.75	0.75	0.50	0.75	69.0	69.0	81.5	79.5	81.5
12	0.95	1.25	0.75	0.50	1.00	99.0	73.0	77.5	60.0	73.0
12	0.95	1.25	1.25	1.00	0.75	70.0	62.5	86.0	68.0	96.5
12	0.95	0.75	0.75	0.75	0.50	52.5	52.0	94.0	94.0	92.0
12	0.65	0.75	0.75	1.00	1.00	57.5	77.5	84.0	78.0	87.5
12	0.95	1.00	1.25	0.50	1.00	88.0	71.5	76.5	72.5	78.0
12	0.95	1.25	0.75	0.50	0.75	98.5	66.5	76.5	70.5	76.0
12	0.95	1.25	1.00	0.75	1.00	100	70.0	75.5	65.0	77.5
12	0.95	1.00	1.25	0.50	0.75	83.0	69.5	78.0	78.0	81.0
12	0.95	1.00	0.75	0.50	1.00	90.0	73.5	78.5	71.0	77.0
12	0.65	1.25	1.00	1.00	0.75	89.0	75.5	77.5	69.5	79.5
12	0.95	0.75	0.75	0.75	1.00	57.5	46.0	91.5	98.0	99.5
12	0.95	1.00	1.00	0.50	0.50	69.5	68.5	79.5	91.5	84.5
12	0.65	1.25	1.00	1.00	1.00	91.0	79.5	80.5	65.5	80.5
12	0.95	1.00	0.75	0.50	0.50	84.5	67.0	80.0	86.5	82.5
12	0.65	1.25	1.25	1.00	0.50	79.0	77.5	78.5	82.0	85.0
12	0.65	1.25	1.25	1.00	0.75	88.5	79.0	81.0	71.5	82.5
12	0.95	0.75	1.00	1.00	0.50	96.5	50.0	100	83.5	76.5
12	0.95	0.75	0.75	0.50	1.00	73.0	72.0	88.5	80.0	94.0
12	0.65	1.25	1.00	1.00	0.50	83.5	78.5	76.5	83.5	85.5
12	0.65	1.25	0.75	1.00	0.50	87.0	80.5	77.5	81.0	83.0
12	0.95	0.75	0.75	0.50	0.50	70.5	71.0	88.0	93.5	87.5
12	0.95	1.25	0.75	0.50	0.50	96.5	66.0	76.5	86.0	86.0
12	0.95	1.25	1.00	0.75	0.75	100	71.0	79.5	72.5	88.5
12	0.95	1.25	1.25	0.75	0.50	67.0	67.0	83.5	96.5	98.0
12	0.65	1.25	0.75	1.00	1.00	99.5	85.5	83.0	61.0	83.5
12	0.65	1.25	0.75	1.00	0.75	98.5	80.0	82.5	69.5	83.5
12	0.95	1.25	0.75	1.00	1.00	100	100	95.5	21.5	98.5
12	0.95	1.00	1.25	0.50	0.50	87.0	73.5	79.0	89.5	87.0
12	0.95	1.25	1.00	1.00	0.75	100	71.5	84.0	68.5	96.0
12	0.95	1.00	0.75	0.50	0.75	92.0	75.5	86.0	81.5	86.0
12	0.65	0.75	0.75	1.00	0.75	86.0	82.5	88.0	82.5	88.0
12	0.95	1.25	1.25	1.00	0.50	72.0	72.0	84.5	99.5	100
12	0.95	0.75	1.00	0.75	0.50	98.0	59.5	100	91.5	88.0
12	0.65	1.00	0.75	0.75	0.75	95.0	85.5	87.5	83.0	88.5
12	0.95	1.25	1.00	0.75	0.50	91.5	75.0	80.0	95.5	98.0
12	0.95	0.75	1.00	1.00	0.75	100	56.5	100	97.5	87.5
12	0.65	0.75	1.25	1.00	0.75	100	78.5	97.0	84.0	82.0
12	0.95	0.75	1.00	0.75	0.75	100	58.5	100	95.5	88.0
12	0.95	1.25	0.75	0.75	0.75	100	85.0	90.5	75.5	92.5
12	0.95	0.75	1.00	0.50	0.75	97.5	77.5	95.0	87.0	87.0
12	0.65	0.75	1.00	1.00	0.75	97.5	83.0	92.0	84.0	88.0
12	0.95	0.75	1.00	0.50	1.00	99.5	82.5	91.5	82.0	91.5
12	0.65	0.75	1.00	1.00	1.00	95.0	83.5	91.0	87.0	91.0
12	0.95	1.25	0.75	0.75	1.00	100	98.5	89.5	69.0	91.0
12	0.95	0.75	1.25	1.00	0.50	100	71.0	100	93.0	85.5
12	0.95	0.75	1.25	1.00	0.75	100	65.0	100	100	85.5
12	0.95	0.75	1.00	0.50	0.50	97.5	78.0	95.5	92.5	87.5
12	0.65	0.75	0.75	1.00	0.50	90.0	90.5	89.0	90.5	91.0
12	0.65	1.00	1.25	0.75	0.75	94.0	87.5	93.0	87.5	89.5
12	0.95	0.75	1.25	0.50	0.75	100	81.0	94.0	92.0	85.5
12	0.65	0.75	1.25	1.00	0.50	96.5	85.5	98.0	86.5	87.0
12	0.65	0.75	1.00	1.00	0.50	92.0	87.0	94.0	91.0	90.0
12	0.65	1.00	1.00	0.75	0.75	91.0	90.5	92.5	90.0	91.0
12	0.95	1.25	1.00	1.00	0.50	97.0	80.0	79.5	99.0	99.5
12	0.65	1.00	1.25	0.75	0.50	92.5	90.0	94.0	91.5	92.0
12	0.95	1.00	1.00	1.00	1.00	60.0	100	100	100	100
12	0.65	1.00	0.75	0.75	0.50	94.0	91.0	91.0	91.5	93.0
12	0.65	0.75	0.75	0.50	1.00	90.5	93.0	93.5	91.5	93.5
12	0.65	0.75	0.75	0.50	0.75	92.0	92.0	93.0	92.5	93.0
12	0.95	0.75	1.25	0.50	1.00	100	88.5	95.0	85.5	94.0
12	0.65	1.00	1.00	0.75	1.00	93.5	91.5	93.5	91.5	93.5
12	0.95	1.25	0.75	1.00	0.75	100	91.5	98.0	74.5	99.5
12	0.95	0.75	1.25	0.75	0.50	100	76.0	100	97.5	90.5
12	0.95	0.75	1.25	0.75	0.75	100	72.5	100	100	91.5
12	0.95	1.00	1.00	0.75	0.75	90.5	87.5	97.0	93.0	97.0
12	0.65	1.00	1.25	0.75	1.00	99.0	91.5	92.5	92.0	91.5
12	0.65	1.00	1.00	1.00	1.00	67.0	100	100	100	100
12	0.95	0.75	1.25	0.50	0.50	100	84.5	96.0	97.0	90.0

Continued on next page

Table 5 – *Continued from previous page*

γ	ρ	β_2	δ_2	β_4	δ_4	πps–reg	STSI–reg	STSI–HT	πps–pos	STSI-pos
12	0.65	0.75	1.00	0.50	0.50	95.5	93.0	93.5	95.5	93.0
12	0.65	0.75	1.25	0.50	1.00	97.5	93.5	93.5	93.5	93.5
12	0.65	0.75	1.00	0.50	1.00	94.5	94.5	95.0	93.0	95.0
12	0.65	0.75	0.75	0.50	0.50	95.0	93.5	94.5	96.5	95.5
12	0.65	0.75	1.25	0.50	0.50	99.0	93.5	93.0	96.0	93.5
12	0.95	1.25	0.75	0.75	0.50	100	90.0	89.5	96.5	99.0
12	0.65	0.75	1.25	1.00	1.00	100	87.0	99.5	98.0	91.0
12	0.65	1.00	0.75	0.75	1.00	99.0	94.5	93.0	94.0	95.0
12	0.65	1.00	1.00	0.75	0.50	95.0	94.5	96.5	94.5	95.5
12	0.65	0.75	1.00	0.50	0.75	95.5	94.5	95.5	95.5	95.5
3	0.65	1.25	1.25	0.50	0.50	94.5	94.5	96.0	96.0	96.0
12	0.65	0.75	1.25	0.50	0.75	99.0	94.0	95.0	95.0	94.5
12	0.95	1.00	1.25	0.75	0.75	100	87.0	99.0	96.0	98.5
12	0.95	1.00	1.00	0.75	1.00	95.0	92.5	98.5	96.5	98.5
12	0.95	1.00	0.75	0.75	0.75	100	93.5	97.5	93.5	97.5
3	0.65	1.25	1.00	0.50	0.75	97.0	96.0	97.0	97.0	97.0
12	0.95	1.00	0.75	0.75	0.50	100	90.5	95.5	99.0	99.5
12	0.95	1.00	1.00	0.75	0.50	93.5	93.0	99.5	99.5	99.5
3	0.65	1.25	0.75	0.50	0.50	98.0	97.0	97.0	97.0	97.0
3	0.65	1.25	1.00	0.50	0.50	97.0	96.5	97.0	98.0	98.0
3	0.65	1.25	0.75	0.50	1.00	100	96.5	97.0	96.5	97.0
3	0.65	1.25	1.25	0.50	0.75	97.0	97.0	97.5	97.5	98.0
3	0.65	1.25	0.75	0.50	0.75	99.5	97.0	97.0	97.0	97.0
12	0.95	1.00	1.25	0.75	1.00	100	92.5	99.5	97.5	100
3	0.65	1.25	1.00	0.50	1.00	98.5	98.0	98.0	98.0	98.0
12	0.95	0.75	1.00	0.75	1.00	100	90.5	100	100	100
12	0.95	1.00	1.25	0.75	0.50	100	91.0	99.5	100	100
12	0.95	0.75	1.00	1.00	1.00	100	91.0	100	100	100
12	0.95	1.00	0.75	0.75	1.00	100	100	98.0	93.0	100
12	0.65	1.00	0.75	1.00	1.00	93.5	100	100	99.5	100
3	0.65	1.00	1.00	0.50	0.50	98.5	98.5	99.0	99.0	99.0
3	0.65	1.00	0.75	0.50	0.75	99.0	98.5	99.0	99.0	99.0
3	0.65	1.00	1.25	0.50	0.50	99.0	98.5	99.0	99.5	99.0
3	0.65	1.25	1.25	0.50	1.00	99.0	99.0	99.0	99.0	99.0
12	0.65	1.00	1.25	1.00	1.00	95.5	100	100	100	100
3	0.65	1.25	1.25	0.75	0.50	99.0	99.0	99.0	99.5	99.5
12	0.95	1.25	0.75	1.00	0.50	100	98.0	98.0	100	100
3	0.95	1.00	1.00	0.50	0.50	98.5	98.0	100	100	100
3	0.95	1.00	1.00	0.50	0.75	98.0	98.5	100	100	100
3	0.65	0.75	0.75	0.50	0.75	98.5	99.0	100	100	100
3	0.65	1.00	0.75	0.50	0.50	99.5	99.5	99.5	99.5	99.5
3	0.65	1.00	0.75	0.50	1.00	99.5	99.5	99.5	99.5	99.5
3	0.65	1.25	1.00	0.75	0.50	99.5	99.5	99.5	99.5	99.5
3	0.65	1.25	1.25	0.75	1.00	99.5	99.5	99.5	99.5	99.5
12	0.95	0.75	1.25	0.75	1.00	100	97.5	100	100	100
3	0.65	0.75	1.00	0.50	0.75	100	99.0	99.5	100	99.5
3	0.65	0.75	1.25	0.50	0.75	100	98.5	100	100	99.5
3	0.65	1.25	1.00	0.75	0.75	100	99.5	99.5	100	99.5
3	0.65	0.75	1.00	0.50	0.50	100	99.5	100	100	99.5
3	0.65	0.75	1.25	0.50	1.00	100	99.0	100	100	100
3	0.95	1.00	1.00	0.50	1.00	99.5	99.5	100	100	100
3	0.65	1.00	1.25	0.50	0.75	100	99.5	99.5	100	100
3	0.95	1.25	1.25	0.50	0.75	99.5	99.5	100	100	100
3	0.95	1.00	0.75	0.50	0.50	100	99.5	100	100	100
3	0.65	1.00	1.00	0.50	0.75	100	99.5	100	100	100
3	0.95	1.00	1.25	0.50	0.75	100	99.5	100	100	100
3	0.95	1.25	1.00	0.50	0.75	100	99.5	100	100	100
12	0.65	0.75	0.75	0.75	1.00	99.5	100	100	100	100
12	0.95	0.75	1.25	1.00	1.00	100	99.5	100	100	100
3	0.65	0.75	0.75	0.50	0.50	100	100	100	100	100
3	0.95	0.75	0.75	0.50	0.50	100	100	100	100	100
3	0.95	0.75	0.75	0.50	0.75	100	100	100	100	100
3	0.65	0.75	0.75	0.50	1.00	100	100	100	100	100
3	0.95	0.75	0.75	0.50	1.00	100	100	100	100	100
3	0.95	0.75	1.00	0.50	0.50	100	100	100	100	100
3	0.95	0.75	1.00	0.50	0.75	100	100	100	100	100
3	0.65	0.75	1.00	0.50	1.00	100	100	100	100	100
3	0.95	0.75	1.00	0.50	1.00	100	100	100	100	100
3	0.65	0.75	1.25	0.50	0.50	100	100	100	100	100
3	0.95	0.75	1.25	0.50	0.50	100	100	100	100	100
3	0.95	0.75	1.25	0.50	0.75	100	100	100	100	100

Continued on next page

Table 5 – *Continued from previous page*

γ	ρ	β_2	δ_2	β_4	δ_4	πps–reg	STSI–reg	STSI–HT	πps–pos	STSI-pos
3	0.95	0.75	1.25	0.50	1.00	100	100	100	100	100
3	0.65	0.75	0.75	0.75	0.50	100	100	100	100	100
3	0.95	0.75	0.75	0.75	0.50	100	100	100	100	100
3	0.65	0.75	0.75	0.75	0.75	100	100	100	100	100
3	0.95	0.75	0.75	0.75	0.75	100	100	100	100	100
3	0.65	0.75	0.75	0.75	1.00	100	100	100	100	100
3	0.95	0.75	0.75	0.75	1.00	100	100	100	100	100
3	0.65	0.75	1.00	0.75	0.50	100	100	100	100	100
3	0.95	0.75	1.00	0.75	0.50	100	100	100	100	100
3	0.65	0.75	1.00	0.75	0.75	100	100	100	100	100
3	0.95	0.75	1.00	0.75	0.75	100	100	100	100	100
3	0.65	0.75	1.00	0.75	1.00	100	100	100	100	100
3	0.95	0.75	1.00	0.75	1.00	100	100	100	100	100
3	0.65	0.75	1.25	0.75	0.50	100	100	100	100	100
3	0.95	0.75	1.25	0.75	0.50	100	100	100	100	100
3	0.65	0.75	1.25	0.75	0.75	100	100	100	100	100
3	0.95	0.75	1.25	0.75	0.75	100	100	100	100	100
3	0.65	0.75	1.25	0.75	1.00	100	100	100	100	100
3	0.95	0.75	1.25	0.75	1.00	100	100	100	100	100
3	0.65	0.75	0.75	1.00	0.50	100	100	100	100	100
3	0.95	0.75	0.75	1.00	0.50	100	100	100	100	100
3	0.65	0.75	0.75	1.00	0.75	100	100	100	100	100
3	0.95	0.75	0.75	1.00	0.75	100	100	100	100	100
3	0.65	0.75	0.75	1.00	1.00	100	100	100	100	100
3	0.95	0.75	0.75	1.00	1.00	100	100	100	100	100
3	0.65	0.75	1.00	1.00	0.50	100	100	100	100	100
3	0.95	0.75	1.00	1.00	0.50	100	100	100	100	100
3	0.65	0.75	1.00	1.00	0.75	100	100	100	100	100
3	0.95	0.75	1.00	1.00	0.75	100	100	100	100	100
3	0.65	0.75	1.00	1.00	1.00	100	100	100	100	100
3	0.95	0.75	1.00	1.00	1.00	100	100	100	100	100
3	0.65	0.75	1.25	1.00	0.50	100	100	100	100	100
3	0.95	0.75	1.25	1.00	0.50	100	100	100	100	100
3	0.65	0.75	1.25	1.00	0.75	100	100	100	100	100
3	0.95	0.75	1.25	1.00	0.75	100	100	100	100	100
3	0.65	0.75	1.25	1.00	1.00	100	100	100	100	100
3	0.95	0.75	1.25	1.00	1.00	100	100	100	100	100
3	0.95	1.00	0.75	0.50	0.75	100	100	100	100	100
3	0.95	1.00	0.75	0.50	1.00	100	100	100	100	100
3	0.65	1.00	1.00	0.50	1.00	100	100	100	100	100
3	0.95	1.00	1.25	0.50	0.50	100	100	100	100	100
3	0.65	1.00	1.25	0.50	1.00	100	100	100	100	100
3	0.95	1.00	1.25	0.50	1.00	100	100	100	100	100
3	0.65	1.00	0.75	0.75	0.50	100	100	100	100	100
3	0.95	1.00	0.75	0.75	0.50	100	100	100	100	100
3	0.65	1.00	0.75	0.75	0.75	100	100	100	100	100
3	0.95	1.00	0.75	0.75	0.75	100	100	100	100	100
3	0.65	1.00	0.75	0.75	1.00	100	100	100	100	100
3	0.95	1.00	0.75	0.75	1.00	100	100	100	100	100
3	0.65	1.00	1.00	0.75	0.50	100	100	100	100	100
3	0.95	1.00	1.00	0.75	0.50	100	100	100	100	100
3	0.65	1.00	1.00	0.75	0.75	100	100	100	100	100
3	0.95	1.00	1.00	0.75	0.75	100	100	100	100	100
3	0.65	1.00	1.00	0.75	1.00	100	100	100	100	100
3	0.95	1.00	1.00	0.75	1.00	100	100	100	100	100
3	0.65	1.00	1.25	0.75	0.50	100	100	100	100	100
3	0.95	1.00	1.25	0.75	0.50	100	100	100	100	100
3	0.65	1.00	1.25	0.75	0.75	100	100	100	100	100
3	0.95	1.00	1.25	0.75	0.75	100	100	100	100	100
3	0.65	1.00	1.25	0.75	1.00	100	100	100	100	100
3	0.95	1.00	1.25	0.75	1.00	100	100	100	100	100
3	0.65	1.00	0.75	1.00	0.50	100	100	100	100	100
3	0.95	1.00	0.75	1.00	0.50	100	100	100	100	100
3	0.65	1.00	0.75	1.00	0.75	100	100	100	100	100
3	0.95	1.00	0.75	1.00	0.75	100	100	100	100	100
3	0.65	1.00	0.75	1.00	1.00	100	100	100	100	100
3	0.95	1.00	0.75	1.00	1.00	100	100	100	100	100
3	0.65	1.00	1.00	1.00	0.50	100	100	100	100	100
3	0.95	1.00	1.00	1.00	0.50	100	100	100	100	100
3	0.65	1.00	1.00	1.00	0.75	100	100	100	100	100
3	0.95	1.00	1.00	1.00	0.75	100	100	100	100	100

Continued on next page

Table 5 – *Continued from previous page*

γ	ρ	β_2	δ_2	β_4	δ_4	πps–reg	STSI–reg	STSI–HT	πps–pos	STSI-pos
3	0.65	1.00	1.00	1.00	1.00	100	100	100	100	100
3	0.95	1.00	1.00	1.00	1.00	100	100	100	100	100
3	0.65	1.00	1.25	1.00	0.50	100	100	100	100	100
3	0.95	1.00	1.25	1.00	0.50	100	100	100	100	100
3	0.65	1.00	1.25	1.00	0.75	100	100	100	100	100
3	0.95	1.00	1.25	1.00	0.75	100	100	100	100	100
3	0.65	1.00	1.25	1.00	1.00	100	100	100	100	100
3	0.95	1.00	1.25	1.00	1.00	100	100	100	100	100
3	0.95	1.25	0.75	0.50	0.50	100	100	100	100	100
3	0.95	1.25	0.75	0.50	0.75	100	100	100	100	100
3	0.95	1.25	0.75	0.50	1.00	100	100	100	100	100
3	0.95	1.25	1.00	0.50	0.50	100	100	100	100	100
3	0.95	1.25	1.00	0.50	1.00	100	100	100	100	100
3	0.95	1.25	1.25	0.50	0.50	100	100	100	100	100
3	0.95	1.25	1.25	0.50	1.00	100	100	100	100	100
3	0.65	1.25	0.75	0.75	0.50	100	100	100	100	100
3	0.95	1.25	0.75	0.75	0.50	100	100	100	100	100
3	0.65	1.25	0.75	0.75	0.75	100	100	100	100	100
3	0.95	1.25	0.75	0.75	0.75	100	100	100	100	100
3	0.65	1.25	0.75	0.75	1.00	100	100	100	100	100
3	0.95	1.25	0.75	0.75	1.00	100	100	100	100	100
3	0.95	1.25	1.00	0.75	0.50	100	100	100	100	100
3	0.95	1.25	1.00	0.75	0.75	100	100	100	100	100
3	0.65	1.25	1.00	0.75	1.00	100	100	100	100	100
3	0.95	1.25	1.00	0.75	1.00	100	100	100	100	100
3	0.95	1.25	1.25	0.75	0.50	100	100	100	100	100
3	0.65	1.25	1.25	0.75	0.75	100	100	100	100	100
3	0.95	1.25	1.25	0.75	0.75	100	100	100	100	100
3	0.95	1.25	1.25	0.75	1.00	100	100	100	100	100
3	0.65	1.25	0.75	1.00	0.50	100	100	100	100	100
3	0.95	1.25	0.75	1.00	0.50	100	100	100	100	100
3	0.65	1.25	0.75	1.00	0.75	100	100	100	100	100
3	0.95	1.25	0.75	1.00	0.75	100	100	100	100	100
3	0.65	1.25	0.75	1.00	1.00	100	100	100	100	100
3	0.95	1.25	0.75	1.00	1.00	100	100	100	100	100
3	0.65	1.25	1.00	1.00	0.50	100	100	100	100	100
3	0.95	1.25	1.00	1.00	0.50	100	100	100	100	100
3	0.65	1.25	1.00	1.00	0.75	100	100	100	100	100
3	0.95	1.25	1.00	1.00	0.75	100	100	100	100	100
3	0.65	1.25	1.00	1.00	1.00	100	100	100	100	100
3	0.95	1.25	1.00	1.00	1.00	100	100	100	100	100
3	0.65	1.25	1.25	1.00	0.50	100	100	100	100	100
3	0.95	1.25	1.25	1.00	0.50	100	100	100	100	100
3	0.65	1.25	1.25	1.00	0.75	100	100	100	100	100
3	0.95	1.25	1.25	1.00	0.75	100	100	100	100	100
3	0.65	1.25	1.25	1.00	1.00	100	100	100	100	100
3	0.95	1.25	1.25	1.00	1.00	100	100	100	100	100
12	0.65	0.75	0.75	0.75	0.50	100	100	100	100	100
12	0.65	0.75	0.75	0.75	0.75	100	100	100	100	100
12	0.65	0.75	1.00	0.75	0.50	100	100	100	100	100
12	0.65	0.75	1.00	0.75	0.75	100	100	100	100	100
12	0.65	0.75	1.00	0.75	1.00	100	100	100	100	100
12	0.65	0.75	1.25	0.75	0.50	100	100	100	100	100
12	0.65	0.75	1.25	0.75	0.75	100	100	100	100	100
12	0.65	0.75	1.25	0.75	1.00	100	100	100	100	100
12	0.65	1.00	0.75	1.00	0.50	100	100	100	100	100
12	0.95	1.00	0.75	1.00	0.50	100	100	100	100	100
12	0.65	1.00	0.75	1.00	0.75	100	100	100	100	100
12	0.95	1.00	0.75	1.00	0.75	100	100	100	100	100
12	0.95	1.00	0.75	1.00	1.00	100	100	100	100	100
12	0.65	1.00	1.00	1.00	0.50	100	100	100	100	100
12	0.95	1.00	1.00	1.00	0.50	100	100	100	100	100
12	0.65	1.00	1.00	1.00	0.75	100	100	100	100	100
12	0.95	1.00	1.00	1.00	0.75	100	100	100	100	100
12	0.65	1.00	1.25	1.00	0.50	100	100	100	100	100
12	0.95	1.00	1.25	1.00	0.50	100	100	100	100	100
12	0.65	1.00	1.25	1.00	0.75	100	100	100	100	100
12	0.95	1.00	1.25	1.00	0.75	100	100	100	100	100
12	0.95	1.00	1.25	1.00	1.00	100	100	100	100	100

The assumption of $N \to \infty$ in result 2 is actually required only for getting closed and nice expressions that can be printed in a paper. Alternatively, one could resort to simulations in order to approximate the anticipated variance. We have developed a program that performs these type of simulations.[1]

5 Conclusions

The strategy that couples πps with the regression estimator is optimal when the superpopulation model exists and some of its parameters are known. This fact was verified by a simulation study in section 3.

Taking into account how strong these assumptions are, it was shown in section 4 that this optimality breaks down when there is a misspecification of the model. Approximations to the anticipated variance of five strategies were obtained for a simple type of misspecification. They were used to verify that πps–reg is not necessarily optimal anymore. In fact its use may lead to variances many times bigger than some other strategies, e.g. STSI–reg, that seem to be more robust.

References

Brewer, K.R.W. (1963). *A Model of Systematic Sampling with Unequal Probabilities.* Australian Journal of Statistics, **5**, 5-13.

Brewer, K.R.W. (2002). *Combined Survey Sampling Inference: Weighing Basu?s Elephants.* London: Arnold.

Cassel, C.M., Särndal, C. E. and Wretman, J. (1977). *Foundations of Inference in Survey Sampling.* New York: Wiley.

Dalenius, T. and Hodges, J.L. (1959) *Minimum variance stratification.* Journal of the American Statistical Association, **54**, 88-101.

Godambe, V.P. (1955). *A unified theory of sampling from finite populations.* Journal of the Royal Statistical Society, Series B **17**, 269-278.

Hanif, M. and Brewer K. R. W. (1980). *Sampling with Unequal Probabilities without Replacement: A Review.* International Statistical Review **48**, 317-335.

Holmberg, A. and Swensson, B. (2001). *On Pareto πps Sampling: Reflections on Unequal Probability Sampling Strategies.* Theory of Stochastic Processes, **7(23)**, 142-155.

Isaki, C.T. and Fuller, W.A. (1982) *Survey design under the regression superpopulation model.* Journal of the American Statistical Association **77**, 89-96.

Kozak, M. and Wieczorkowski, R. (2005). *πps Sampling versus Stratified Sampling ? Comparison of Efficiency in Agricultural Surveys.* Statistics in Transition, **7**, 5-12.

Lanke, J. (1973). *On UMV-estimators in Survey Sampling.* Metrika **20**, 196 202.

[1]For information about the program, you can visit the webpage of the Department of Statistics at Stockholm University (https://www.statistics.su.se) or email the author.

Rosén, B. (1997). *On sampling with probability proportional to size.* Journal of statistical planning and inference **62**, 159-191.

Rosén, B. (2000a). *Generalized Regression Estimation and Pareto πps.* R&D Report 2000:5. Statistics Sweden.

Rosén, B. (2000b). *On inclusion probabilities for order πps sampling.* Journal of statistical planning and inference **90**, 117-143.

Särndal, C.E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling.* Springer.

Tillé, Y. (2006). *Sampling algorithms.* Springer.

Wright, R.L. (1983). *Finite Population Sampling with Multivariate Auxiliary Information.* Journal of the American Statistical Association, **78**, 879 884.

Appendix. Proof of results

Proof of Result 1. The following expectations are required in the proof of the result

$$\mathrm{E}_\xi Y_k = \mathrm{E}_\xi \left[\beta_0 + \beta_1 x_k^{\beta_2} + \epsilon_k\right] = \beta_0 + \beta_1 x_k^{\beta_2} \tag{14}$$

$$\mathrm{E}_\xi Y_k^2 = \mathrm{E}_\xi \left[\left(\beta_0 + \beta_1 x_k^{\beta_2} + \epsilon_k\right)^2\right] = \mathrm{E}_\xi \left[\left(\beta_0 + \beta_1 x_k^{\beta_2}\right)^2 + 2\left(\beta_0 + \beta_1 x_k^{\beta_2}\right)\epsilon_k + \epsilon_k^2\right] = \left(\beta_0 + \beta_1 x_k^{\beta_2}\right)^2 + 2\left(\beta_0 + \beta_1 x_k^{\beta_2}\right)\mathrm{E}_\xi\left[\epsilon_k\right] + \mathrm{E}_\xi\left[\epsilon_k^2\right] = \left(\beta_0 + \beta_1 x_k^{\beta_2}\right)^2 + \beta_3^2 x_k^{2\beta_4} \tag{15}$$

$\mathrm{E}_\xi \overline{Y}$, $\mathrm{E}_\xi \overline{Y^2}$ and $\mathrm{E}_\xi \overline{xY}$ are obtained using (14) and (15),

$$\mathrm{E}_\xi \overline{Y} = \mathrm{E}_\xi \left[\frac{1}{N}\sum_U Y_k\right] = \frac{1}{N}\sum_U \mathrm{E}_\xi Y_k = \frac{1}{N}\sum_U (\beta_0 + \beta_1 x_k^{\beta_2}) = \beta_0 + \beta_1 \overline{x^{\beta_2}} \tag{16}$$

$$\mathrm{E}_\xi \overline{Y^2} = \mathrm{E}_\xi \left[\frac{1}{N}\sum_U Y_k^2\right] = \frac{1}{N}\sum_U \left(\left(\beta_0 + \beta_1 x_k^{\beta_2}\right)^2 + \beta_3^2 x_k^{2\beta_4}\right) = \beta_0^2 + 2\beta_0\beta_1\overline{x^{\beta_2}} + \beta_1^2\overline{x^{2\beta_2}} + \beta_3^2\overline{x^{2\beta_4}} \tag{17}$$

$$\mathrm{E}_\xi \overline{xY} = \mathrm{E}_\xi \left[\frac{1}{N}\sum_U x_k Y_k\right] = \frac{1}{N}\sum_U x_k \mathrm{E}_\xi Y_k = \frac{1}{N}\sum_U x_k(\beta_0 + \beta_1 x_k^{\beta_2}) = \beta_0\overline{x} + \beta_1\overline{x^{\beta_2+1}} \tag{18}$$

Now, using (16), (17) and (18) we get

$$\mathrm{E}_\xi\left[\overline{xY} - \overline{x}\,\overline{Y}\right] = (\beta_0\overline{x} + \beta_1\overline{x^{\beta_2+1}}) - \overline{x}(\beta_0 + \beta_1\overline{x^{\beta_2}}) = \beta_1(\overline{x^{\beta_2+1}} - \overline{x}\overline{x^{\beta_2}}) = \beta_1 S_{1,\beta_2} \tag{19}$$

$$\mathrm{E}_\xi\left[\overline{Y^2} - \overline{Y}^2\right] = (\beta_0^2 + 2\beta_0\beta_1\overline{x^{\beta_2}} + \beta_1^2\overline{x^{2\beta_2}} + \beta_3^2\overline{x^{2\beta_4}}) - (\beta_0 + \beta_1\overline{x^{\beta_2}})^2 = \beta_1^2\left(\overline{x^{2\beta_2}} - \overline{x^{\beta_2}}^2\right) + \beta_3^2\overline{x^{2\beta_4}} = \beta_1^2 S_{\beta_2,\beta_2} + \beta_3^2\overline{x^{2\beta_4}} \tag{20}$$

Using (19) and (20), we obtain an approximation to the correlation coefficient, $R_{x,y}$,

$$R_{x,y}^2 = \frac{(\overline{xy} - \overline{x}\overline{y})^2}{(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2)} \approx \frac{\mathrm{E}_\xi^2\left[\overline{xY} - \overline{x}\,\overline{Y}\right]}{\mathrm{E}_\xi\left[(\overline{x^2} - \overline{x}^2)\left(\overline{Y^2} - \overline{Y}^2\right)\right]} = \frac{\beta_1^2 S_{1,\beta_2}^2}{S_{1,1}(\beta_1^2 S_{\beta_2,\beta_2} + \beta_3^2\overline{x^{2\beta_4}})} \tag{21}$$

Solving (21) for β_3^2 we get (8), as desired. □

Proof of Result 2. Parts 1 and 2. Some results are required before we obtain (9) and (10). If the model ξ is true but the reg-estimator with $\boldsymbol{x}_k = (1, x_k^{\delta_2})$ is fitted, we have

$$Y_k = B_0 + B_1 x_k^{\delta_2} + E_k \qquad \text{but also} \qquad Y_k = \beta_0 + \beta_1 x_k^{\beta_2} + \epsilon_k$$

where $B_0 = \frac{t_y}{N} - B_1 \frac{t_{x^{\delta_2}}}{N}$ and $B_1 = \frac{N t_{x^{\delta_2} y} - t_{x^{\delta_2}} t_y}{N t_{x^{2\delta_2}} - t^2_{x^{\delta_2}}}$. Then

$$\mathrm{E}_\xi B_1 = \mathrm{E}_\xi \left[\frac{N t_{x^{\delta_2} y} - t_{x^{\delta_2}} t_y}{N t_{x^{2\delta_2}} - t^2_{x^{\delta_2}}} \right] = \frac{1}{N t_{x^{2\delta_2}} - t^2_{x^{\delta_2}}} \left(N \sum_U x_k^{\delta_2} \mathrm{E}_\xi Y_k - t_{x^{\delta_2}} \sum_U E_\xi Y_k \right) =$$
$$\frac{1}{N t_{x^{2\delta_2}} - t^2_{x^{\delta_2}}} \left(N \sum_U x_k^{\delta_2} (\beta_0 + \beta_1 x_k^{\beta_2}) - t_{x^{\delta_2}} \sum_U (\beta_0 + \beta_1 x_k^{\beta_2}) \right) =$$
$$\frac{1}{N t_{x^{2\delta_2}} - t^2_{x^{\delta_2}}} \left(N\beta_0 \sum_U x_k^{\delta_2} + N\beta_1 \sum_U x_k^{\beta_2+\delta_2} - N\beta_0 \sum_U x_k^{\delta_2} - \beta_1 \sum_U x_k^{\delta_2} \sum_U x_k^{\beta_2} \right) =$$
$$\beta_1 \frac{\overline{x^{\beta_2+\delta_2}} - \overline{x^{\beta_2}}\,\overline{x^{\delta_2}}}{\overline{x^{2\delta_2}} - \overline{x^{\delta_2}}^2} = \beta_1 \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \quad (22)$$

$$\mathrm{E}_\xi B_0 = \mathrm{E}_\xi \left[\frac{t_y}{N} - B_1 \frac{t_{x^{\delta_2}}}{N} \right] = \frac{1}{N} \sum_U \mathrm{E}_\xi Y_k - \overline{x^{\delta_2}} \mathrm{E}_\xi B_1 =$$
$$\frac{1}{N} \sum_U (\beta_0 + \beta_1 x_k^{\beta_2}) - \overline{x^{\delta_2}} \beta_1 \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} = \beta_0 + \beta_1 \overline{x^{\beta_2}} - \beta_1 \overline{x^{\delta_2}} \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} =$$
$$\beta_0 - \beta_1 \left(\overline{x^{\delta_2}} \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} - \overline{x^{\beta_2}} \right) \quad (23)$$

Using (22) and (23) we obtain $\mathrm{E}_\xi E_k$ and an approximation to $\mathrm{E}_\xi E_k^2$,

$$\mathrm{E}_\xi E_k = \mathrm{E}_\xi \left[\beta_0 - B_0 + \beta_1 x_k^{\beta_2} - B_1 x_k^{\delta_2} + \epsilon_k \right] = \beta_0 - \mathrm{E}_\xi B_0 + \beta_1 x_k^{\beta_2} - \mathrm{E}_\xi B_1 x_k^{\delta_2} + \mathrm{E}_\xi \epsilon_k =$$
$$\beta_0 - \beta_0 + \beta_1 \left(\overline{x^{\delta_2}} \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} - \overline{x^{\beta_2}} \right) + \beta_1 x_k^{\beta_2} - \beta_1 \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} x_k^{\delta_2} + 0 =$$
$$\beta_1 \left(\left(x_k^{\beta_2} - \overline{x^{\beta_2}} \right) - \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(x_k^{\delta_2} - \overline{x^{\delta_2}} \right) \right) \quad (24)$$

$$\mathrm{E}_\xi E_k^2 = \mathrm{E}_\xi \left[\left(\beta_0 - B_0 + \beta_1 x_k^{\beta_2} - B_1 x_k^{\delta_2} + \epsilon_k \right)^2 \right] \approx$$
$$\left(\beta_0 - B_0 + \beta_1 x_k^{\beta_2} - B_1 x_k^{\delta_2} \right)^2 + 2 \left(\beta_0 - B_0 + \beta_1 x_k^{\beta_2} - B_1 x_k^{\delta_2} \right) \mathrm{E}_\xi \epsilon_k + \mathrm{E}_\xi \epsilon_k^2 \approx$$
$$\left(\beta_0 - \beta_0 - \beta_1 \left(\overline{x^{\delta_2}} \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} + \overline{x^{\beta_2}} \right) + \beta_1 x_k^{\beta_2} - \beta_1 \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} x_k^{\delta_2} \right)^2 + \beta_3^2 x_k^{2\beta_4} =$$
$$\beta_1^2 \left(\left(x_k^{\beta_2} - \overline{x^{\beta_2}} \right) - \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(x_k^{\delta_2} - \overline{x^{\delta_2}} \right) \right)^2 + \beta_3^2 x_k^{2\beta_4} \quad (25)$$

The approximations in (25) are obtained by assuming $N \to \infty$, so that B_0 and B_1 have essentially no variance under the model. $\mathrm{E}_\xi \overline{E_h}$ and an approximation to $\mathrm{E}_\xi \overline{E_h^2}$

are obtained using (24) and (25),

$$\mathrm{E}_\xi \overline{E_h} = \mathrm{E}_\xi \left[\frac{1}{N_h} \sum_{U_h} E_k \right] = \frac{1}{N_h} \sum_{U_h} \mathrm{E}_\xi E_k = \frac{1}{N_h} \sum_{U_h} \left(\beta_1 \left(\left(x_k^{\beta_2} - \overline{x^{\beta_2}} \right) - \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(x_k^{\delta_2} - \overline{x^{\delta_2}} \right) \right) \right) = \beta_1 \left(\left(\overline{x_h^{\beta_2}} - \overline{x^{\beta_2}} \right) - \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(\overline{x_h^{\delta_2}} - \overline{x^{\delta_2}} \right) \right) \quad (26)$$

$$\mathrm{E}_\xi \overline{E_h^2} = \mathrm{E}_\xi \left[\frac{1}{N_h} \sum_{U_h} E_k^2 \right] = \frac{1}{N_h} \sum_{U_h} \mathrm{E}_\xi E_k^2 \approx \frac{1}{N_h} \sum_{U_h} \left(\beta_1^2 \left(\left(x_k^{\beta_2} - \overline{x^{\beta_2}} \right) - \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(x_k^{\delta_2} - \overline{x^{\delta_2}} \right) \right)^2 + \beta_3^2 x_k^{2\beta_4} \right) =$$
$$\frac{1}{N_h} \sum_{U_h} \left(\beta_1^2 \left(\left(x_k^{\beta_2} - \overline{x^{\beta_2}} \right)^2 - 2\frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(x_k^{\beta_2} - \overline{x^{\beta_2}} \right) \left(x_k^{\delta_2} - \overline{x^{\delta_2}} \right) + \frac{S_{\beta_2,\delta_2}^2}{S_{\delta_2,\delta_2}^2} \left(x_k^{\delta_2} - \overline{x^{\delta_2}} \right)^2 \right) + \beta_3^2 x_k^{2\beta_4} \right) =$$
$$\beta_1^2 \left(\left(\overline{x_h^{2\beta_2}} - 2\overline{x^{\beta_2}}\,\overline{x_h^{\beta_2}} + \overline{x^{\beta_2}}^2 \right) - 2\frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(\overline{x_h^{\beta_2+\delta_2}} - \overline{x^{\delta_2}}\,\overline{x_h^{\beta_2}} - \overline{x^{\beta_2}}\,\overline{x_h^{\delta_2}} + \overline{x^{\beta_2}}\,\overline{x^{\delta_2}} \right) + \frac{S_{\beta_2,\delta_2}^2}{S_{\delta_2,\delta_2}^2} \left(\overline{x_h^{2\delta_2}} - 2\overline{x^{\delta_2}}\,\overline{x_h^{\delta_2}} + \overline{x^{\delta_2}}^2 \right) \right) + \beta_3^2 \overline{x_h^{2\beta_4}} \quad (27)$$

Assuming that the population is large enough, we approximate the variance of πps sampling by the variance of with-replacement pps, i.e.

$$\mathrm{E}_\xi \mathrm{V}_{\pi\mathrm{ps}} \left(\hat{t}_{\mathrm{reg}} \right) \approx \mathrm{E}_\xi \left[\frac{1}{n} \sum_U p_k \left(\frac{E_k}{p_k} - t_E \right)^2 \right] \approx \frac{1}{n} \sum_U \frac{\mathrm{E}_\xi E_k^2}{p_k} \quad (28)$$

with $p_k = \frac{x_k^{\delta 4}}{t_{x^{\delta_4}}}$. The last approximation is obtained by using $t_E \approx 0$. Using (25) into (28) we get

$$\frac{1}{n} \sum_U \frac{\mathrm{E}_\xi E_k^2}{p_k} \approx \frac{1}{n} \sum_U \frac{\beta_1^2 \left(\left(x_k^{\beta_2} - \overline{x^{\beta_2}} \right) - \frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(x_k^{\delta_2} - \overline{x^{\delta_2}} \right) \right)^2 + \beta_3^2 x_k^{2\beta_4}}{\frac{x_k^{\delta 4}}{t_{x^{\delta_4}}}} =$$
$$\frac{N^2}{n} \overline{x^{\delta_4}} \left(\beta_1^2 \left(\left(\overline{x^{2\beta_2-\delta_4}} - 2\overline{x^{\beta_2}}\,\overline{x^{\beta_2-\delta_4}} + \overline{x^{\beta_2}}^2\,\overline{x^{-\delta_4}} \right) - 2\frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(\overline{x^{\beta_2+\delta_2-\delta_4}} - \overline{x^{\delta_2}}\,\overline{x^{\beta_2-\delta_4}} - \overline{x^{\beta_2}}\,\overline{x^{\delta_2-\delta_4}} + \overline{x^{\beta_2}}\,\overline{x^{\delta_2}}\,\overline{x^{-\delta_4}} \right) + \frac{S_{\beta_2,\delta_2}^2}{S_{\delta_2,\delta_2}^2} \left(\overline{x^{2\delta_2-\delta_4}} - 2\overline{x^{\delta_2}}\,\overline{x^{\delta_2-\delta_4}} + \overline{x^{\delta_2}}^2\,\overline{x^{-\delta_4}} \right) \right) + \beta_3^2 \overline{x^{2\beta_4-\delta_4}} \right) =$$
$$\frac{N^2}{n} \overline{x^{\delta_4}} \left(\beta_1^2 \left(\left(S_{\beta_2,\beta_2-\delta_4} - \overline{x^{\beta_2}} S_{\beta_2,-\delta_4} \right) - 2\frac{S_{\beta_2,\delta_2}}{S_{\delta_2,\delta_2}} \left(S_{\beta_2,\delta_2-\delta_4} - \overline{x^{\delta_2}} S_{\beta_2,-\delta_4} \right) + \frac{S_{\beta_2,\delta_2}^2}{S_{\delta_2,\delta_2}^2} \left(S_{\delta_2,\delta_2-\delta_4} - \overline{x^{\delta_2}} S_{\delta_2,-\delta_4} \right) \right) + \beta_3^2 \overline{x^{2\beta_4-\delta_4}} \right)$$

Replacing β_3^2 by (8) we obtain (9).

Regarding the variance of STSI we have

$$\mathrm{E}_\xi \mathrm{V}_{\mathrm{STSI}}\left(\hat{t}_{\mathrm{reg}}\right) = \mathrm{E}_\xi\left[\sum_{h=1}^{H}\frac{N_h^2}{n_h}\left(1-\frac{n_h}{N_h}\right)\frac{N_h}{N_h-1}\left(\overline{E_h^2}-\overline{E_h}^2\right)\right] \approx$$

$$\mathrm{E}_\xi\left[\sum_{h=1}^{H}\frac{N_h^2}{n_h}\left(\overline{E_h^2}-\overline{E_h}^2\right)\right] \approx \sum_{h=1}^{H}\frac{N_h^2}{n_h}\left(\mathrm{E}_\xi\overline{E_h^2}-\mathrm{E}_\xi\overline{E_h}^2\right) \quad (29)$$

The approximations are obtained assuming that $N \to \infty$, so that, in the first case, the with-replacement variance can be used instead of its without-replacement counterpart, and, in the second case, the variance of $\overline{E_h}$ is essentially zero. Using (26), (27) and (8) into (29) we obtain (10), as desired. □

Proof of Result 2. Part 3. Some results are required before we obtain (11). Using 14 and 15 we get

$$\mathrm{E}_\xi\overline{Y_h} = \mathrm{E}_\xi\left[\frac{1}{N_h}\sum_{U_h}Y_k\right] = \frac{1}{N_h}\sum_{U_h}\mathrm{E}_\xi Y_k = \frac{1}{N_h}\sum_{U_h}\left(\beta_0+\beta_1 x_k^{\beta_2}\right) = \beta_0+\beta_1\overline{x_h^{\beta_2}} \quad (30)$$

$$\mathrm{E}_\xi\overline{Y_h^2} = \mathrm{E}_\xi\left[\frac{1}{N_h}\sum_{U_h}Y_k^2\right] = \frac{1}{N_h}\sum_{U_h}\mathrm{E}_\xi Y_k^2 = \frac{1}{N_h}\sum_{U_h}\left(\beta_0^2+2\beta_0\beta_1\overline{x_h^{\beta_2}}+\beta_1^2\overline{x_h^{2\beta_2}}+\beta_3^2\overline{x_h}^{2\beta_4}\right) \quad (31)$$

The anticipated variance of strategy 3 is

$$\mathrm{E}_\xi \mathrm{V}_{\mathrm{STSI}}\left(\hat{t}_{\mathrm{HT}}\right) = \mathrm{E}_\xi\left[\sum_{h=1}^{H}\frac{N_h^2}{n_h}\left(1-\frac{n_h}{N_h}\right)\frac{N_h}{N_h-1}\left(\overline{Y_h^2}-\overline{Y_h}^2\right)\right] \approx$$

$$\mathrm{E}_\xi\left[\sum_{h=1}^{H}\frac{N_h^2}{n_h}\left(\overline{Y_h^2}-\overline{Y_h}^2\right)\right] \approx \sum_{h=1}^{H}\frac{N_h^2}{n_h}\left(\mathrm{E}_\xi\overline{Y_h^2}-\mathrm{E}_\xi\overline{Y_h}^2\right)$$

Using (30), (31) and (8) we obtain (11) as desired. □

Proofs of Result 2. Parts 4 and 5. Some results are needed before we obtain (12) and (13). If the model ξ is true, but the pos-estimator as defined for strategies 4 and 5 in section 2.2 is fitted, we have

$$Y_k = B_g + E_k \qquad \text{but also} \qquad Y_k = \beta_0 + \beta_1 x_k^{\beta_2} + \epsilon_k$$

where $B_g = \frac{1}{N_g}\sum_{U_g} Y_k$. Then

$$\mathrm{E}_\xi B_g = \mathrm{E}_\xi\left[\frac{1}{N_g}\sum_{U_g}Y_k\right] = \frac{1}{N_g}\sum_{U_g}\mathrm{E}_\xi Y_k = \frac{1}{N_g}\sum_{U_g}\left(\beta_0+\beta_1 x_k^{\beta_2}\right) = \beta_0+\beta_1\overline{x_g^{\beta_2}} \quad (32)$$

Using (32) we obtain $\mathrm{E}_\xi E_k$ and and approximation to $\mathrm{E}_\xi E_k^2$,

$$\mathrm{E}_\xi E_k = \mathrm{E}_\xi \left[\beta_0 + \beta_1 x_k^{\beta_2} - B_g + \epsilon_k\right] = \beta_0 + \beta_1 x_k^{\beta_2} - \mathrm{E}_\xi B_g + \mathrm{E}_\xi \epsilon_k = \beta_0 + \beta_1 x_k^{\beta_2} - \beta_0 - \beta_1 \overline{x_g^{\beta_2}} = \beta_1 \left(x_k^{\beta_2} - \overline{x_g^{\beta_2}}\right) \quad (33)$$

$$\mathrm{E}_\xi E_k^2 = \mathrm{E}_\xi \left[\left(\beta_0 + \beta_1 x_k^{\beta_2} - B_g + \epsilon_k\right)^2\right] \approx \left(\beta_0 + \beta_1 x_k^{\beta_2} - B_g\right)^2 + 2\left(\beta_0 + \beta_1 x_k^{\beta_2} - B_g\right) \mathrm{E}_\xi \epsilon_k + \mathrm{E}_\xi \epsilon_k^2 \approx \left(\beta_0 + \beta_1 x_k^{\beta_2} - \beta_0 - \beta_1 \overline{x_g^{\beta_2}}\right)^2 + \beta_3^2 x_k^{2\beta_4} = \beta_1^2 \left(x_k^{\beta_2} - \overline{x_g^{\beta_2}}\right)^2 + \beta_3^2 x_k^{2\beta_4} \quad (34)$$

$\mathrm{E}_\xi \overline{E_h}$ and an approximation to $\mathrm{E}_\xi \overline{E_h^2}$ are obtained using (33) and (34),

$$\mathrm{E}_\xi \overline{E_h} = \frac{1}{N_h} \sum_{U_h} \mathrm{E}_\xi E_k = \frac{1}{N_h} \sum_{U_h} \beta_1 \left(x_k^{\beta_2} - \overline{x_g^{\beta_2}}\right) = \frac{\beta_1}{N_h} \sum_{g=1}^{G} N_{hg} \left(\overline{x_{hg}^{\beta_2}} - \overline{x_g^{\beta_2}}\right) \quad (35)$$

$$\mathrm{E}_\xi \overline{E_h^2} = \frac{1}{N_h} \sum_{U_h} \mathrm{E}_\xi E_k^2 \approx \frac{1}{N_h} \sum_{U_h} \left(\beta_1^2 \left(x_k^{\beta_2} - \overline{x_g^{\beta_2}}\right)^2 + \beta_3^2 x_k^{2\beta_4}\right) = \frac{1}{N_h} \left(\beta_1^2 \left(\sum_{g=1}^{G} N_{hg} \overline{x_{hg}^{2\beta_2}} - 2 \sum_{g=1}^{G} N_{hg} \overline{x_g^{\beta_2} x_{hg}^{\beta_2}} + \sum_{g=1}^{G} N_{hg} \overline{x_g^{\beta_2}}^2\right) + \beta_3^2 N_h \overline{x_h^{2\beta_4}}\right) = \frac{1}{N_h} \left(\beta_1^2 \sum_{g=1}^{G} N_{hg} \left(\overline{x_{hg}^{2\beta_2}} - 2\overline{x_g^{\beta_2} x_{hg}^{\beta_2}} + \overline{x_g^{\beta_2}}^2\right) + \beta_3^2 N_h \overline{x_h^{2\beta_4}}\right) \quad (36)$$

Assuming that the population is large enough, the anticipated variance of strategy 4 is approximated by (28) with $\mathrm{E}_\xi E_k^2$ approximated by (34), and we obtain

$$\frac{1}{n} \sum_{U} \frac{\mathrm{E}_\xi E_k^2}{p_k} \approx \frac{1}{n} \sum_{U} \frac{\beta_1^2 \left(x_k^{\beta_2} - \overline{x_g^{\beta_2}}\right)^2 + \beta_3^2 x_k^{2\beta_4}}{\frac{x_k^{\delta_4}}{t_{x^{\delta_4}}}} = \frac{N\overline{x^{\delta_4}}}{n} \left(\beta_1^2 \left(\sum_{U} x_k^{2\beta_2 - \delta_4} - 2 \sum_{U} \overline{x_g^{\beta_2}} x_k^{\beta_2 - \delta_4} + \sum_{U} \overline{x_g^{\beta_2}}^2 x_k^{-\delta_4}\right) + \beta_3^2 \sum_{U} x_k^{2\beta_4 - \delta_4}\right) = \frac{N\overline{x^{\delta_4}}}{n} \left(\beta_1^2 \left(\sum_{g=1}^{G} N_g \overline{x_g^{2\beta_2 - \delta_4}} - 2 \sum_{g=1}^{G} N_g \overline{x_g^{\beta_2}}\, \overline{x_g^{\beta_2 - \delta_4}} + \sum_{g=1}^{G} N_g \overline{x_g^{\beta_2}}^2 \overline{x_g^{-\delta_4}}\right) + \beta_3^2 \sum_{U} x_k^{2\beta_4 - \delta_4}\right) \quad (37)$$

Using (8) into (37), we obtain the desired result (12).

Finally, (13) is obtained in an analogous way, using (35), (36) and (8) into

$$\mathrm{E}_\xi V_{\mathrm{STSI}}\left(\hat{t}_{\mathrm{pos}}\right) \approx \sum_{h=1}^{H} \frac{N_h^2}{n_h} \left(\mathrm{E}_\xi \overline{E_h^2} - \mathrm{E}_\xi \overline{E_h}^2\right)$$

□