Review report PWC matrices 2016/2040

Gunnar Flötteröd, VTI

November 5, 2019

1 Introduction and summary of reviewed report

This document is composed as follows. The present section summarizes the reviewed report. Section 2 provides a summary assessment of the deployed method and its result. Section 3 discusses several major points in greater depth. A possible further development path is sketched in Section 4.

The present document is complemented by the following previously submitted material:

- Detailed comments in the form of two annotated draft project reports (one per review seminar).
- Major discussion items of the review seminars in the form of two sets of slides (one per review seminar).
- Preliminary "Suggestions for estimation and prediction of PWC matrices" (enclosed under this title in the first review report).

The reviewed report consists of ten sections. Section 1 describes the PWC matrix structure and outlines the matrix estimation process. Section 2 describes preparatory input data transformations. Section 3 presents the preparation of historical commodity value time series. Section 4 documents the estimation of a gravity-type commodity flow model. Section 5 describes the preparation of data defining marginal conditions for the estimation of the 2016 PWC matrices. Section 6 describes the estimation of the 2016 PWC matrices through a balancing of the output of Section 4's gravity model such that the marginal conditions of Section 5 are reproduced. Section 7 compares the obtained 2016 PWC matrices to available 2012 PWC matrices. Section 8 describes the preparation of data defining marginal conditions for the 2040 PWC matrices, including a time series model for the prediction of commodity values. Section 9 presents the predicted 2040 PWC matrices. Section 10 summarizes elements of the report and enumerates deviations from a previously established estimation process.[1]

2 Summary assessment of method and result

PWC (production/warehouse/consumption) matrices represent a fixed demand for goods transportation. The PWC matrices produced in the reviewed project serve as input to the national freight transportation model Samgods. There is one matrix for each of in total 16 commodity types. Each matrix is composed

[1] C. Anderstig, M. Berglund, H. Edwards, M. Sundberg (2015). PWC Matrices: new method and updated Base Matrices. Final project report.

of blocks for domestic, import, export and transit flows. The matrices do not differentiate between flows from production to a warehouse and from a warehouse to consumption. Seasonal variability is not modeled.

The PWC matrices are produced in a multi-step process that combines a large number of different data sources into matrix point estimates and predictions per commodity. The overall structure of the process was established prior to the reviewed project; modifications of the original approach are enumerated in Section 10 of the reviewed report. The method comprises three major elements.

1. Estimation of past commodity values (Section 3) and their prediction based on a time-series model (Section 8); these values are used to transform goods and monetary flows.

2. Estimation of a gravity model describing inter-zonal goods flows (Section 4), using primarily data from Sweden's commodity flow survey (VFU).

3. Alignment of the gravity model's predictions with current resp. predicted boundary conditions through a matrix balancing procedure (Section 5 and 6 for reproduction of 2016 matrices; Section 8 and 9 for prediction of 2040 matrices).

This structure alone is reasonably clear. Its concrete implementation, however, is not. The sheer number of interacting data manipulation steps in combination with in parts counter-intuitive or unclear mathematical operations makes it for the reviewer impossible to assess the quality of the produced matrices. (Details can be found in the previously submitted material.) Two main reasons for this difficulty can be named:

- The reviewed report ties together many data sources, which differ in their scope, aggregation level, time stamp, completeness, quality, unit.

- The reviewed report does not clearly separate data, model specification, and model estimation resp. evaluation/prediction.

There appears to be room for improvement in both regards, even though this would likely exceed what can be achieved within the reviewed project. In terms of the applicability of the produced matrices, the reviewer can only recommend to not uncritically use these as point estimates of a status quo or as predictors of a possible future. It is necessary to keep in mind the (in parts major) uncertainties arising in the estimation/prediction process.

In terms of formal requirements as enumerated in Trafikverket's review request, the following can be stated.

1. *Analys och översyn av nuvarande metoder för framtagning av basårsmatriser, varuvärdesprognoser, utrikesvaruhandelsprognoser, transitprognoser och prognosårsmatriser.*

   The method description comprises (i) references to relevant elements of the predecessor report and (ii) a description of deviations from that report. A review of alternative approaches, e.g. from other Nordic or European countries, is not included.

2. *Ny skattning av gravitationsmodeller för generering av matriser m.h.t. en ny varugruppsindelning i Samgodsmodellen, ny varuflödesundersökning VFU2016, senast tillgänglig övrig statistik, m.m.*

   Included.

3. *Implementering av framtagna metoder i verktyg/modell för en automatiserad framtagning av matriser.*

   It is unclear how well automated the process is. The described process uses several different programming and scripting languages; to what extent their interplay is (or can be) automated is not explained.

4. *Tillämpning av de metoder och verktyg/modeller som utvecklats och skattats genom:*

   (a) *framtagning av basårsmatriser (m.h.a. så kallade randvillkor, inklusive ev. komplettering av kända, större flöden, samt nedbrytning till företagsnivå),*
   Included.

   (b) *framtagning av varuvärdesprognos (inklusive underlag för att beräkna så kallade godstidsvärden för basår och prognosår),*
   A time-series based commodity value model is included. How this can be used to compute "godstidsvärden" is not clear, given that this term does not appear in the reviewed document.

   (c) *framtagning av utrikesvaruhandelsprognos,*
   Import/export are included.

   (d) *framtagning av transitprognos,*
   Transit is included.

   (e) *framtagning av prognosårsmatriser (m. h. a. så kallade randvillkor, inklusive nedbrytning till företagsnivå).*
   Prognosis matrices are included. A "nedbrytning till företagsnivå" in terms of a disaggregation of flows by business size is not explicitly described in Sections 8 and 9 but may have been conducted without explicit mention in the reviewed report.

5. *Kvalitetsgranskning. Analys av framtagna basårs- och prognosårsmatriser genom kvalitetsgranskning och testning i Samgodsmodellen, validering mot tillgänglig statistik om transportarbete m.m, inklusive eventuella justeringar.*

   Included in Section 6 for 2016 and, to the extent possible, in Section 9 for 2040. (The comparison between 2012 and 2016 PWC matrices in Sec. 7 is, as the authors themselves state, not meaningful.)

6. *Uppdraget skall även utmynna i en slutrapport avseende alla uppnådda resultat. Rapportutkast levereras uppdragsgivaren senast en månad innan uppdragets avslut. Denna aktivitet gäller även eventuell option.*

   A report is available.

7. *Eventuell option avser arbetet med känslighetsanalyser.*

   A limited sensitivity analysis is included: singular flows are studied in Sec. 4.4.2; Section 6.2 makes statements with respect to the limited explanatory power of commodity value point estimates but does not provide interpretable numerical evidence.

3 Major comments and recommendations

Trafikverket requested this report to be complemented with recommendations for future improvements. The present section serves this purpose, by discussing possibilities to overcome what the reviewer perceives as three main weaknesses of the reviewed approach.

3.1 Estimation and prediction of commodity values

The imprecision of the estimated commodity values appears to be a major weakness. The following is stated in Section 6.2 of the reviewed report:

> *"Känslighetsanalysen inkludera även varuvärdena i modellestimeringarna. Det visade väsentligen att varuvärdena har mycket låga förklaringsvärden, få blev signifikanta och resultaten ligger långt från de skattningar som erhålls från varuvärdesmodellen. Vi kom till motsvarande slutsats i arbetet med PWC-matriserna för år 2012. Varuvärdena uppvisar helt enkelt en mycket stor variation för de heterogena varor som ingår i våra stora aggregat av varugrupper."*

In brief, point estimates of commodity values are inadequate to represent a possibly high within-commodity value variability. This appears reasonable, but further data and/or analysis in support of this hypothesis are not included in the reviewed report. It appears necessary to better understand the available data. The commodity flow survey (VFU) contains many *individual* transactions (with both shipment size and value) that would allow for such an investigation. Examples of simple but potentially insightful analysis steps follow below.

- Plot a histogram of commodity values per commodity. How large is the variability? Are the values centered around a mean or is there more structure, e.g. multiple peaks that would suggest that several distinct groups of goods with different values are combined? Is there a systematic difference between PW and WC values?
- Plot a map (of Sweden) showing mean values per commodity and commune (one map with communes as senders, one with communes as receivers). Same for variability. Is there a spatial (e.g. north/south or east/west or coast/inland) pattern? Is there a dependence between nearby communes? Are there communes that stick out? Is there a systematic difference between PW and WC values?

The insights obtained in such an analysis could then guide the development of a commodity value *model*, i.e. a representation of commodity values as functions of observable attributes (e.g. features of the sending/receiving commune, attributes of the shipment, ...); this is currently not the case (Section 3 of the reviewed report). As touched upon in the first review seminar, such a model could be estimated jointly with the gravity flow model, i.e. with a limited additional implementation effort (details on this in the supplementery material of the first review report). This approach could replace the rather unreliable selection of domestic commodity values from import/export values (as described in the beginning of Sec. 6.1 and in Table 6.1 for 2016 estimation, and in Sec. 8.3.1 for 2040 prediction).

The reviewed report describes the year-to-year commodity value dynamics with a time series model that does not include exogenous explanatory variables (Section 8.4). It may be possible to identify economic indicators that can be used as explanatory variables in a further developed version of this model. This may allow to predict commodity values based on predictions of these indicators.

3.2 Separation of specification and model estimation

The reviewer stressed already in the first review seminar the importance of separating model specification and model estimation/evaluation. Not doing so leads to a difficult-to-understand modeling process and hardly interpretable results. In the second review seminar, the reviewer illustrated these problems by

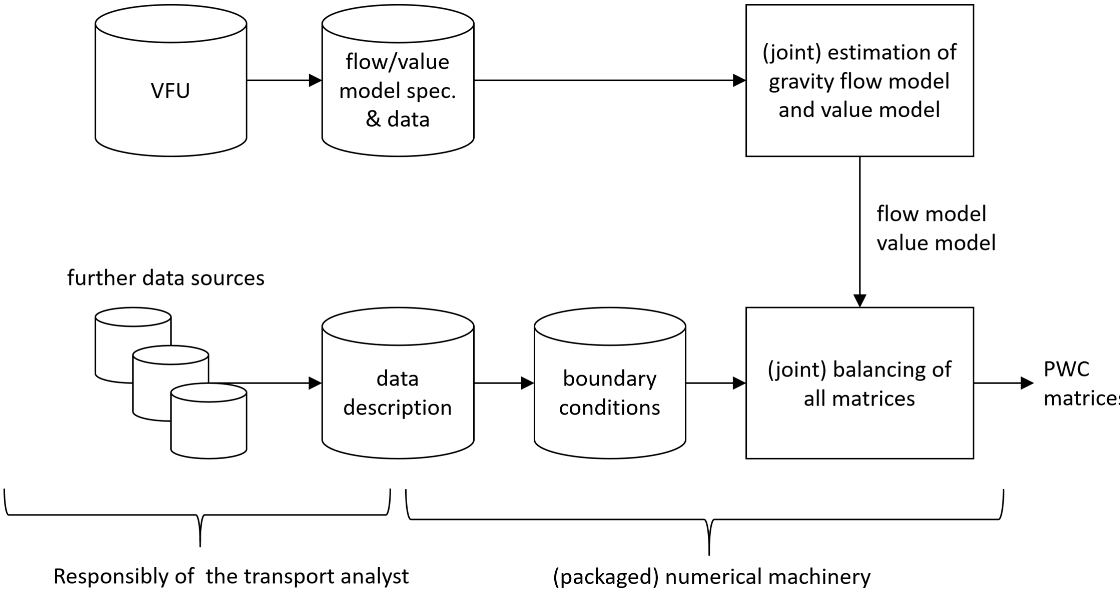

Figure 1: PWC estimation workflow separation

reverse-engineering two moments of the modeling process into simple but interpretable mathematical models.

It appears possible to at least reduce problems of this type. The remainder of this section describes an idealized and certainly incomplete picture of how this could look; a concrete development path leading to such a configuration would require further analysis and specification efforts. A very much simplified outline is provided in Figure 1.

The upper half of Figure 1 shows the present setup of estimating the commodity flow gravity model, only that it now includes a commodity value model, as suggested in Section 3.1. Importantly in terms of workflow, the transport analyst focuses on the encoding of data and the specification of a model structure; the mathematics and numerics of the model estimation are packaged and automated.

The lower half of Figure 1 now shows a similar separation between problem specification and solution for the matrix balancing step. Most importantly, the transport analyst is relieved of the burden to specify all boundary conditions in the same unit, to resolve inconsistencies, or to fill in missing data. The analyst focuses on annotating and submitting the available data for analysis. Examples of such annotations are: (i) units (e.g. SEK or tons), (ii) to which matrix entries the data set applies (e.g. only import or export or domestic), (iii) possibly value ranges if the data source is uncertain or ambiguous. It is then up to the packaged numerical machinery to translate this information into boundary conditions for (an improved version of) the matrix balancing problem:

Some data sources may contain information that applies to multiple commodities. Without further processing, such data hence leads to boundary conditions for more than one commodity matrix. As explained in the supplementary material of the first review report, this can be accounted for by (i) formulating combined boundary conditions for all commodity matrices and then (ii) jointly balancing all commodity matrices. The format according to which the transport analyst specifies the available data complies with a supplementary *measurement model*, which is described in the following section.

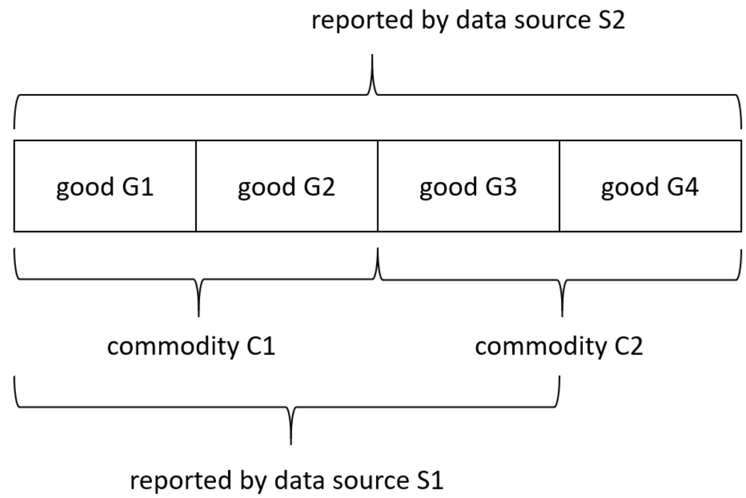

Figure 2: Measurement model example

3.3 Explicit measurement model

The available data sources may differ in many regards: scope, aggregation level, time stamp, completeness, quality, unit. The reviewed report describes an approach where data is transformed and combined in a process-based manner. This process is both error-prone and difficult to interpret. The previous Section 3.2 advocates a workflow where the transport analyst focuses on a description of the available data, but leaves its (numerical) processing to a packaged algorithm. This requires a unified data representation through which information at all available levels of (dis-)aggregation can be represented; this is well in line with an observation from the second review seminar, where an (internal) data representation with more than 16 commodities was considered. The benefits of such a representation are subsequently illustrated through a simple example; see also Figure 2.

Assume there are four types of goods (G_1 to G_4), which are aggregated into two commodities C_1 and C_2 as follows:

$$C_1 = G_1 + G_2 \quad (1)$$
$$C_2 = G_3 + G_4. \quad (2)$$

There further are two data sources S_1 and S_2 available, which, however, do not exactly report the commodities of interest:

$$S_1 = G_1 + G_2 + G_3 \quad (3)$$
$$S_2 = G_1 + G_2 + G_3 + G_4. \quad (4)$$

That is, S_1 reports C_1 and a part of C_2, and S_2 reports the sum of C_1 and C_2.

The reviewed report describes an approach that would in the present example correspond to estimating (possibly based on secondary data sources or some other reasoning) the magnitude of the overlap between data sources and commodities, and to feeding the resulting estimators of C_1 and C_2 into the further matrix estimation process.

Alternatively, one could for as long as possible abstain from manipulating the available data and merely document its properties. In the present case, this would amount to annotating the unmodified values S_1 and S_2 with supplementary information in the form of the *measurement model* (3),(4). This would correspond to the "data description" shown in Figure 1. – If the subsequent matrix balancing process then happens to succeed in computing a reliable estimator for G_3, the following re-arrangement of the measurement model (1)-(4) would infer the desired commodity values:

$$C_1 = S_1 - G_3 \quad (5)$$

$$C_2 = S_2 - S_1 + G_3. \quad (6)$$

This example may, due to its simplicity, appear anecdotal. Given the real setting, in which a large (and hardly manageable) number of data points is submitted for analysis, the transport analyst will not be able to assess *a priori* the best possible combination of all data sources, which is, however, possible for a suitably designed measurement model and estimation algorithm. Very much in the way the gravity model estimation algorithm provides supplementary statistics (e.g. parameter significance values), the machinery for formulating boundary conditions could automatically identify (and possibly even correct) inconsistencies or data gaps.

4 A possible development path

The estimation and prediction of commodity values appears to be one of the weakest elements of the current PWC matrix estimation approach. However, a structured and documented analysis of this problem appears to have not yet been undertaken. This is hence recommended as a first improvement step. The first half of Section 3.1 enumerates several simple yet potentially insightful analyses of the available VFU data that could serve as a starting point. Depending on the result of this analysis, further improvement steps could then be undertaken. If the within-commodity value variability turns out to be strongly coupled to observable quantities, a sensible next step would be the development of a commodity value model along the lines of the second half of Section 3.1.

If a streamlining of the overall matrix estimation process is desirable, the proposals of Sections 3.2 and 3.3 could be considered. Arguably, this would be a less incremental undertaking. One could first study concrete formulations for the joint balancing of all commodity matrices with a joint set of boundary conditions, i.e. allowing for boundary conditions that apply to more than one commodity (e.g. to their sum). This is more an algorithmic exercise, the feasibility of which is discussed in the supplementary material of the first review report. Given that a convincing approach is found, one could then turn to the specification of a practical data description format, jointly with an underlying measurement model through which this data description can be automatically transformed into boundary conditions for the joint matrix balancing problem.