

**Reconciling Bilateral Trade Statistics in the Presence of
Re-exports via Third Countries:
The Case of China, Hong Kong and Their Major Trading Partners**

Zhi Wang

United States International Trade Commission

Mark Gehlhar

United States Department of Agriculture

Shunli Yao

China Center for Economic Research, Peking University

May 2007

ABSTRACT

This paper develops a mathematical programming model to simultaneously estimate re-export markups and reconcile bilateral trade statistics between China, Hong Kong, and their trading partners. The model is applied to GTAP sector level trade flows to resolve discrepant reporting in an efficient manner. Adjustments in trade flows are based upon particular bilateral trade routines reliability as well as statistical reporter's reliability information. The program is implemented in GAMS and retains many desirable theoretical and empirical properties. Estimates are used for generating trade flows and markups for Hong Kong's re-exports used in the forthcoming version 7 GTAP database. The model's flexibility has potential for expanded use in other regions where re-exports and associated markup cause discrepant trade flows.

JEL Classification Numbers: F1, C61, C81

The views expressed in this paper are solely those of the authors, and are not meant to represent in any way the views of the institutions with which they are affiliated. The authors thank helpful suggestions from K.C. Fung, Michael Ferrantino, Judy Dean, Bill Powers and the participants in the Conference on Discrepancies in US-China Trade Statistics, National Bureau of Statistics, Beijing, China, October 24, 2006. The authors are also grateful to Thomas Hertel at GTAP Center of Purdue University and three anonymous referees for their valuable in depth technical comments. However, all remaining errors are solely the authors' responsibility.

I. Introduction

It has long been known that bilateral trade statistics reported by importing and exporting countries are unlikely to be the same, and in fact they vary greatly from each other for a wide variety of reasons. Economists and statistical agencies around the globe working on reconciling bilateral trade have adopted methods for choosing either the importer's or exporter's data, or some weighted average of the two, as more reliable (e.g. Gehlhar (1996), for the GTAP model, and the documentation for Statistics Canada's *World Trade Analyzer*). However, the standard methods for data reconciliation have generally not worked well for China and its major trading partners because the intermediary role of Hong Kong in China's external trade.

A large share of China's trade with the world passes through Hong Kong, while current reporting practices in China and their trading partners do not fully reflect this fact. This is in part because traders often do not know the final destinations when goods leave China. In these cases, they are recorded as exports to Hong Kong by the Chinese Customs. For this reason, Chinese Customs statistics show that Hong Kong is one of China's largest export destinations, behind the United States but at par with the EU 15 countries in recent years. In fact, Hong Kong re-exports most its imports from China to other countries. On the other hand, the US Customs treats all goods from China, directly or indirectly through Hong Kong, as Chinese imports, including the value added to the goods by Hong Kong middlemen. As a result, discrepancies in the official data on the bilateral trade arise, and its increasingly large magnitude has not only caused concerns among policy makers in the two countries, but has also motivated quite a few studies to reconcile the conflicting official trade statistics between China and its major trading partners.

Key components of those studies on trade data reconciliation in the case of China and Hong Kong include estimation of Hong Kong re-export markups, which are key information but not part of the Hong Kong official trade statistics, as well as the cif/fob ratios which are a traditional concept in explaining discrepancies in official trade statistics reported by exporting and importing countries. On the re-export markup estimation, there are two threads in the literature. One is based on detailed trade data, including studies by the Joint Commission on Commerce and Trade (JCCT) (1995), using solely Hong Kong trade data, and by Feenstra et al (1998, 1999), using both China and Hong Kong trade data; and the other is based on surveys conducted by the Hong Kong Census and Statistical Department (HKCSD) and published in various issues of the *Hong Kong Monthly Digest of Statistics*, and interviews reported in Fung (1996) and Fung and Lau (1998). Among these estimates, Feenstra et al (1998, 1999) are able to produce origin- and destination-specific markups to reconcile various aggregate estimates reported in JCCT (1995), HKCSD, Fung (1996) and Fung and Lau (1998). Subsequent studies on the reconciliation of recent Chinese trade flow with the US, Canada and 69 trading partners follow either the survey and interview approach (Fung and Lau, 2001, 2003; Fung, Lau and Xiong, 2006; Schindler and Beckett, 2005), or combine it with the JCCT approach (Bohatyretz and Santarossa, 2005).

On the estimation of cif/fob ratios, almost all above mentioned studies use an *ad hoc* and one size-for-all estimate, though differing in value across studies. Fung, Lau and Xiong (2006)

even undertake to convert the fas to fob value for the US exports and to include the services trade in their China-US trade data reconciliation.

Those studies attempt to use a large amount of trade statistics to estimate the Hong Kong re-export markups, to include into data adjustment a wide range of factors contributing to the discrepancies and to identify behaviors of traders that may lead to mis-invoicing of China and Hong Kong's trade statistics. They constitute valuable contributions to improving our understanding of the complicated issues. In these studies, however, estimation of the Hong Kong re-export markups, estimation of the cif-fob ratios and adjustment of Chinese trade flow with a country does not depend on each other, nor do they depend on the estimation and data adjustment with other countries. Therefore, the theoretically intrinsic global consistency for the world trade data can not be found in the adjusted bilateral trade flow in those studies. In addition, those studies never fully utilize all official trade statistics from China, Hong Kong and their trading partners simultaneously. This calls for a new approach to trade data reconciliation, which is the very motivation of this paper.

The paper has two goals. First, it develops and implements a formal model to simultaneously estimate Hong Kong re-export markups and reconcile the Chinese and Hong Kong trade statistics in a globally consistent optimization framework. Second, it applies the model to 2004 bilateral world trade data to produce a set of trade estimates for the next version of the Global Trade Analysis Project (GTAP) database (version 7). To enhance its empirical quality, this paper also draws more detailed information on trade related shipping costs to estimate cif/fob ratios and the most updated research on the estimation of Hong Kong re-export markups as inputs into the mathematical programming model.

The paper is organized as follows. Section two specifies the optimization framework and discusses its theoretical and empirical properties. Section three outlines the major steps to implement the model with real world trade statistics, including the preparation of initial fob/cif ratio and Hong Kong's re-export markup estimates, aggregation issues and the choice and estimation of reliability weights for major variables in the model. Modeling results are presented and compared with the initial estimates in section four. The paper concludes with a discussion on limitations of the study and directions of future research.

II. The Mathematical Programming Model

2.1 General Assumptions and Mathematical Notations

Consider China and Hong Kong both engage in bilateral trade with N partner countries and each other on M commodities for time period T . Hong Kong is the only entrepot between China and the N partner countries engaging re-export activities to transship both China's and its N partner countries' exports to each others. Hong Kong earns a markup by conducting such activities. This is basically the difference between the price Hong Kong buys goods and what it sells the same goods for. All partner countries except one report their exports to and imports from China and Hong Kong. China and Hong Kong also report their exports to and imports from all their partner countries and trade flows between them. In addition, Hong Kong reports the origin and destination of all commodities it re-exports bound for and coming from China and other partner countries. The markup from such activities is unreported, thus it must be estimated. Assuming all reporting countries, including China, can correctly identify the country of origin of their imports, either the imports are directly from the partners or indirectly from Hong Kong. Reporters however can not determine the final destination when exports leave their ports (Schindler and Beckett, 2005). The notation used to describe the reported trade statistics and their relationships are as follows (expressed in annual values in this paper):

DX_{it}^{sr} = Direct exports of commodity i from country s to country r at time t . For s equals Hong Kong, it is domestic exports including earnings from re-export that commodity. For r equals Hong Kong, it is partner countries' exports remain in Hong Kong

RX_{it}^{sr} = Indirect exports of commodity i via Hong Kong from origin country s to destination country r at time t , including Hong Kong's re-export earnings

TX_{it}^{sr} = Total or actual exports of commodity i from country s to country r at time t . For s equals Hong Kong, it is its domestic exports plus re-exports

DM_{it}^{sr} = Direct imports of commodity i by country r from country s at time t . For r equals Hong Kong, it is imports for domestic use, for s equals Hong Kong it is partner's imports originated from Hong Kong

TM_{it}^{sr} = Total imports of commodity i by country s from country r at time t

RXM_{it}^{sr} = Hong Kong markup earnings by re-export commodity i originated from country s to final destination country r at time t

WEX_{it}^s = Total exports of commodity i to the world by country s at time t , including both direct and indirect exports to all countries

WMX_{it}^r = Total imports of commodity i from the world by country r at time t , including both direct and indirect imports from all countries

XER_{it}^r = Statistical discrepancy of commodity i in China and Hong Kong's east bound trade with partner country r at time t

MER_{it}^r = Statistical discrepancy of commodity i in China and Hong Kong's west bound trade with partner country r at time t

cif_{it}^{sr} = fob/cif ratio for commodity i shipped from country s to country r at time t . It is a fixed parameter in the model.

Indices i defined over commodity set $I \in \{1, 2, \dots, M\}$, indices s and r defined over country set $W \in \{1, 2, \dots, N, CH, HK\}$. All the trade flow variables have directions. The first superscripts always indicate the source country and the second always refer to destination countries. For exports (DX and TX), source country are the reporter, while for imports (DM and TM), destination country are the reporter. Exports are valued at fob basis and imports are valued at cif basis.

Using notations defined above, the following 16 accounting identities describe the relationship among bilateral trade flow statistics reported by China, Hong Kong and their partner countries.

2.2 Eastbound flows: China and Hong Kong exports, partner imports

For all $r \in \{1, 2, \dots, N\}$ and all $s \in \{1, 2, \dots, N, CH\}$:

$$TX_{it}^{CH,r} + DX_{it}^{HK,r} + XER_{it}^r = cif_{it}^{CH,r} TM_{it}^{CH,r} + cif_{it}^{HK,r} DM_{it}^{HK,r} \quad (1)$$

Equation (1) states that the sum of any particular partner's imports of China and Hong Kong originated products after fob/cif adjustment should equal to the sum of China's total exports and Hong Kong's domestic exports to that partner, plus a statistical discrepancy.

$$TX_{it}^{CH,r} = cif_{it}^{CH,HK} (RX_{it}^{CH,r} - RXM_{it}^{CH,r}) + DX_{it}^{CH,r} \quad (2)$$

Equation (2) defines that China's total exports to a particular partner equal China's direct exports plus Hong Kong's re-exports for China to that partner minus Hong Kong's re-export makeup adjusted by China-Hong Kong fob/cif ratio.

$$DX_{it}^{HK,r} = TX_{it}^{HK,r} - \sum_s (RX_{it}^{sr} - RXM_{it}^{sr}) \quad (3)$$

Equation (3) defines that Hong Kong's domestic exports to a particular partner equals to its total exports to that partner minus its re-exports for all other countries to the particular partner and plus its markup earnings from re-exports¹.

¹ The definition of this variable is different with domestic exports statistics published by Hong Kong authority, which is total exports minus re-exports without adjustment for markup.

$$DM_{it}^{HK,r} = TM_{it}^{HK,r} - \frac{\sum_r (RX_{it}^{sr} - RXM_{it}^{sr})}{cif_{it}^{HK,r}} \quad (4)$$

Equation (4) indicates partner's imports of Hong Kong's domestic products equals partners' total imports from Hong Kong minus Hong Kong's re-exports to the partner from all sources adjusted by Hong Kong re-export markup and fob/cif ratio from Hong Kong to the partner.

$$DM_{it}^{CH,r} = TM_{it}^{CH,r} - \frac{cif_{it}^{CH,HK} (RX_{it}^{CH,r} - RXM_{it}^{CH,r})}{cif_{it}^{CH,r}} \quad (5)$$

Equation (5) indicates that partner's direct imports from China equals its total imports from China minus Hong Kong's re-exports for China to that partner adjusted by Hong Kong's re-exports markup and China to Hong Kong and China to partner cif/fob ratios.

2.3 Westbound flows: China and Hong Kong imports, partner exports

For all $s \in \{1, 2, \dots, N\}$ and all $r \in \{1, 2, \dots, N, CH\}$:

$$cif_{it}^{s,CH} DM_{it}^{s,CH} + cif_{it}^{s,HK} TM_{it}^{s,HK} + MER_{it}^s = DX_{it}^{s,CH} + TX_{it}^{s,HK} \quad (6)$$

Equation (6) states that the sum of China and Hong Kong's total imports of products originated from any particular partner after fob/cif adjustment minus Hong Kong re-exports to China for the partner countries should equal to the sum of that partner's direct exports to China and its total exports to Hong Kong plus a statistical discrepancy.

$$DM_{it}^{s,HK} = TM_{it}^{s,HK} - (\sum_r RX_{it}^{sr} - \sum_r RXM_{it}^{sr}) \quad (7)$$

Equation (7) requires Hong Kong's domestic use of imports plus its re-exports for a particular partner minus re-exports markup equals Hong Kong's total imports from that partner country.

$$DM_{it}^{s,CH} = TM_{it}^{s,CH} - \frac{cif_{it}^{s,HK} (RX_{it}^{s,CH} - RXM_{it}^{s,CH})}{cif_{it}^{s,CH}} \quad (8)$$

Equation (8) defines that China's direct imports from a partner equals China's total imports from that partner minus Hong Kong's re-exports to China for that partner adjusted by Hong Kong's re-export earnings and partner to Hong Kong and partner to China fob/cif ratios.

$$TX_{it}^{s,CH} = DX_{it}^{s,CH} + cif_{it}^{s,HK} (RX_{it}^{s,CH} - RXM_{it}^{s,CH}) \quad (9)$$

Equation (9) reveals that partner's total exports to China equals partner's direct exports to China plus Hong Kong's re-exports to China for that partner adjust by Hong Kong's re-export markup and partner to Hong Kong's fob/cif ratio.

$$DX_{it}^{s,HK} = TX_{it}^{s,HK} - cif_{it}^{s,HK} \sum_r (RX_{it}^{sr} - RXM_{it}^{sr}) \quad (10)$$

Equation (10) defines that a partner's exports to Hong Kong for Hong Kong domestic use equals its total export to Hong Kong minus its re-exports via Hong Kong to all destinations adjust by Hong Kong's re-export markup and partner to Hong Kong's fob/cif ratio.

2.4 China-Hong Kong bilateral trade

$$TX_{it}^{CH,HK} = DX_{it}^{CH,HK} - cif_{it}^{CH,HK} \sum_r (RX_{it}^{CH,r} - RXM_{it}^{CH,r}) \quad (11)$$

$$DX_{it}^{HK,CH} = TX_{it}^{HK,CH} - \sum_r (RX_{it}^{s,CH} - RXM_{it}^{s,CH}) \quad (12)$$

Equation (11) defines that China's actual exports to Hong Kong for Hong Kong domestic use equals its direct exports to Hong Kong minus Hong Kong's re-exports for China to all other trading partners adjusted by Hong Kong re-exports markup and China to Hong Kong fob/cif ratio. Equation (12) defines that Hong Kong's domestic exports to China equals its total exports to China minus its re-exports to China from all other partners adjust by its markup earnings.

$$DM_{it}^{CH,HK} = TM_{it}^{CH,HK} - \sum_r (RX_{it}^{CH,r} - RXM_{it}^{CH,r}) \quad (13)$$

$$TM_{it}^{HK,CH} = DM_{it}^{HK,CH} + \frac{\sum_r (RX_{it}^{s,CH} - RXM_{it}^{s,CH})}{cif_{it}^{HK,CH}} \quad (14)$$

Equations (13) defines Hong Kong's imports from China for domestic use equals its total imports from China minus its re-exports for China to all destinations adjusted by its markup earnings. While equation (14) defines China's total imports from Hong Kong equals its imports of goods with Hong Kong origin plus Hong Kong's re-exports to China from all sources adjusted by re-exports markup and Hong Kong to China fob/cif ratio.

2.5 Global balance and objective function

For all $r \in \{1, 2, \dots, N, CH, HK\}$:

$$\sum_r TX_{it}^{CH,r} + \sum_r DX_{it}^{HK,r} = WEX_{it}^{HK} + WEX_{it}^{CH} \quad (15)$$

$$\sum_r TM_{it}^{s,CH} + \sum_s \sum_r (RX_{it}^{sr} - RXM_{it}^{sr}) + \sum_r DM_{it}^{s,HK} = WMX_{it}^{HK} + WMX_{it}^{CH} \quad (16)$$

Equation (15) states that the sum of after adjustment actual exports from China and Hong Kong to all its partners should still equal to the sum of their reported total exports to the world. It means the adjustment made by the model do not change the total exports to the world reported by China and Hong Kong, it merely estimate Hong Kong's re-export markup and rearrange the destinations of China's exports. Equation (16) states that China and Hong Kong's imports and Hong Kong's re-exports minus re-exports markup after adjustment should still equal to the sum of China and Hong Kong's total imports from the world. The adjustments made by the model only change the markup estimates and rearrange the sources of China and Hong Kong's imports.

In addition, China and Hong Kong's total exports to and imports from the world should satisfy following conditions: total world exports by all trading countries equals total world imports after fob/cif adjustment.

$$\sum_s WEX_{it}^s = \sum_r cif_{it}^r WMX_{it}^r \quad (17)$$

Given above clearly defined accounting relationship among trade flow statistics, what remains to mathematically formulate the reconciliation problem in an optimization framework is the construction of a criteria for changing the reported statistics to conform the know linear accounting constraints. Either a cross-entropy (Harrigan & Buchanan, 1984, Golan et al., 1994) or a quadratic objective penalty function can be specified. We choose to use a quadratic function as follows for computation efficiency reasons²:

$$\begin{aligned} \text{Min} \quad S = \frac{1}{2} \left\{ \begin{aligned} & \sum_t \sum_{i \in M} \sum_{s \in W} \sum_{r \in W} \frac{(DX_{it}^{sr} - DX_{it}^{sr0})^2}{wdx_{it}^{sr}} + \sum_t \sum_{i \in ms \in W} \sum_{r \in W} \frac{(DM_{it}^{sr} - DM_{it}^{sr0})^2}{wdm_{it}^{sr}} \\ & + \sum_t \sum_{i \in M} \sum_{s \in W} \sum_{r \in W} \frac{(TX_{it}^{sr} - TX_{it}^{sr0})^2}{wtx_{it}^{sr}} + \sum_t \sum_{i \in ms \in W} \sum_{r \in W} \frac{(TM_{it}^{sr} - TM_{it}^{sr0})^2}{wtm_{it}^{sr}} \\ & + \sum_t \sum_{i \in ms \in W} \sum_{r \in W} \frac{(RXM_{it}^{sr} - RXM_{it}^{sr0})^2}{wrxm_{it}^{sr}} + \sum_t \sum_{i \in M} (xer_{it}^2 + mer_{it}^2) \end{aligned} \right\} \quad (18) \end{aligned}$$

Where variables with a 0 in the end denote initial estimates for that variable, and an additional “w” before the variable in lower case indicates the reliability measure for that variable.

In short, the reconciliation problem is to modify a given set of bilateral trade flow statistics with equation (18) as objective function and equations (1) to (16) as constraints.

² The quadratic function has a numerical advantage in implementing the model. It is easier to solve than the entropy function in very large models because they can use software specifically designed for quadratic programming. As showed by Canning and Wang (2005), the quadratic function is equivalent to the entropy function in the neighborhood of initial estimates, under a properly selected weighing scheme.

2.6 Properties of the reconciliation model

There are desirable analytical properties of the optimization model specified above. Firstly, the estimates of markups and trade flow adjustments are made in a consistent simultaneous manner. The model re-directs sources and destinations of China's and Hong Kong's exports and imports, estimates Hong Kong's re-export markup, allocates statistical discrepancies to trade flows among China, Hong Kong and their trading partners, and adjust bilateral trade balance for China and all its partners simultaneously. In doing so it imposes global consistency to the adjusted trade flow data, which is a necessary condition for any world trade data set can be used for global general equilibrium trade policy analysis.

Secondly, the model is a nonlinear programming problem subject to only linear constraints, and its solutions could represent a broad range of estimators for linear statistical model by choosing different reliability weights. In other words, there are different statistical interpretations could be given to the model when different weights are used in the objective function. For instance, if weights are all equal to one, the solution of the model gives a constrained least squares estimator. If initial estimates are taken as the weights, the solution of the model gives a weighted constrained least square estimator, which is identical to the Friedlander-solution, and a good approximation of the RAS solution. If the weights are proportional to the variances of the initial estimates and the initial estimates are statistically independent, the solution of the model yields best linear unbiased estimates of the true unknown matrix (Byron, 1978), which is identical to the Generalized Least Squares estimator if the weights are equal to the variance of initial estimates (Stone, 1984, Ploeg, 1984). Furthermore, as noted by Stone et al. (1942) and proven by Weale (1985), in cases where the error distributions of the initial estimates are normal, the solution also satisfies the maximum likelihood criteria.

Thirdly, by understanding the model's solution as estimators of a underlying statistical model and assuming the initial estimates are unbiased estimates of the true unknown values, in all but the trivial case, the adjusted estimates from the model solution will always better approximate the unknown true values than do the associated initial estimates, as proven by Harrigan (1990). This is because adding valid constraints or further restricting the feasible set through the narrowing of interval constraints cannot move the adjusted estimates away from the true values, unless the additional constraints are non-binding (have no information value). The optimization process has the effect of reducing, or at least not increasing, the variance of the initial estimates. This desired property is simple to show by using matrix notation. Define $\mathbf{W}$ as the variance matrix of initial estimates $\bar{\mathbf{D}}$, $\mathbf{R}$ as the coefficient matrix of all linear constraints. The least squares solution (equivalent to the solution of the quadratic programming model described above) to the problem of adjusting $\bar{\mathbf{D}}$ to $\mathbf{D}$ that satisfies the linear constraint, $\mathbf{R} \cdot \mathbf{D} = \mathbf{0}$ can be written as:

$$\mathbf{D} = (\mathbf{I} - \mathbf{WR}^T(\mathbf{RWR}^T)^{-1}\mathbf{R}) \bar{\mathbf{D}} \quad (19)$$

Thus,

$$\text{var}(\mathbf{D}) = (\mathbf{I} - \mathbf{WR}^T(\mathbf{RWR}^T)^{-1}\mathbf{R})\mathbf{W} = \mathbf{W} - \mathbf{WR}^T(\mathbf{RWR}^T)^{-1}\mathbf{R}\mathbf{W} \quad (20)$$

Since $\mathbf{WR}^T(\mathbf{RWR}^T)^{-1}\mathbf{R}\mathbf{W}$ is a positive semi-definite matrix, the variance of adjusted estimates will always be less, or at least not greater than the variance of the initial estimates as long as $\mathbf{R}\bullet\mathbf{D} = \mathbf{0}$ holds³. This is the fundamental reason why such a reconciliation framework will provide better adjusted trade statistics. Imposing equation (1) to (16) will definitely improve, or at least not worsen the initial statistics, since we are sure from international economics that those constraints are consistency requirements and must be true for any well defined trade statistics.

Finally, the choice of weights ($wdx_{it}^{sr}, wtx_{it}^{sr}, wdm_{it}^{sr}, wtm_{it}^{sr}, wrxm_{it}^{sr}$) in the objective function has very important impacts on the model solution. The model uses these weights to determine by how much an initial estimate may be changed. For instance, using the initial trade statistics as weights has the advantage that each entry of the trade flow data is adjusted in proportion to its magnitude in order to satisfy those consistency constraints. The variables can not change signs and the larger the trade flows, the more adjustment takes place. However, the adjustment relates directly to the size of the initial trade statistics, and does not force the unreliable trade data to absorb the bulk of the required adjustment. Furthermore, only under the assumptions: (1) the initial estimates for different trade flows are statistically independent, and (2) each error variance is proportional to the corresponding initial estimates, this commonly used weighing scheme (underlying RAS) can obtain best unbiased estimates, while those assumptions often not hold for international trade data. Therefore, the efficiency of the model will be improved if the error structure of the initial trade statistics is available. Because using such a weighting scheme makes the adjustment independent of the size of the initial trade data. The larger the variance, the smaller its contribution to the objective function, and hence the lesser the penalty for each adjusted trade statistics to move away from their initial value (only the relative, not the absolute size of the variance affects the solution). A small variance of the initial trade statistics indicates, other things equal, it is a very reliable reported data and thus should not change by much, whilst a large variance of the initiate estimates indicates an unreliable report data and may be adjusted considerably, i.e. adjust the trade data in an unreliable reported route more than the reliable report one, thus providing an attractive but feasible weighing mechanism in reconciliation of bilateral trade statistics.

Advantages of such an optimization framework in adjusting international trade statistics are also significant from empirical perspective. Firstly, it offers great convenience and details. Hong Kong's re-export markup rate, each country's re-exports via Hong Kong as percent of the country's total exports and imports, and adjusted bilateral balance of trade among China, Hong Kong and their partner countries by each covered commodity are all part of the model solution.

Secondly, it provides considerable flexibility. It permits a wider variety and volume of information to be brought into the reconciliation process. For example, the ability of introducing upper and/or lower bounds is one of the flexibilities not offered by commonly used scaling procedures such as RAS. Therefore, it is very easy to restrict the value of the adjusted trade statistics to nonnegative in the reconciliation process. This is a very desirable

³ Details of the derivation of equation (19) and (20) can be found in classic textbook of econometrics, such as of *Econometric Methods, second edition* by Johnston, pp 157-158.

property of adjusting bilateral trade flow data. It is also very flexible regarding to the required known information. For example, it allows the possibility that some of the bilateral trade statistics are missing and the total exports and imports by China and Hong Kong to the world do not known with certainty. In the real world, missing bilateral trade is common and a country's total exports or imports generally lie in a range. By incorporation of associated terms similar to bilateral trade variables in the objective function to penalize solution deviations from the world totals from statistical sources, allows reconciliation of these world totals together with bilateral trade flows.

Finally, various relative measures of the reliability of the initial data can be easily included in the reconciliation process, because the choice of values for those reliability weights in the objective function is also very flexible. As noted before, these weights should reflect the relative reliability of the original trade statistics. The interpretation is straightforward. Statistics with higher reliability should be changed less than statistics with a lower reliability, thus the best available information can always be used to insure that statistics reported by reliable trade routes or reporters are not perturb by the reconciliation process as much as statistics reported by unreliable trade routes or reporters.

III. Linking the Model with Trade Statistics

There are several key steps in implementing the optimization model specified above. First, all variables in the model need to be correctly linked with officially reported statistics; second, Hong Kong's markup earning from its re-exports and all bilateral fob to cif margins have to be computed independently or estimated based on information from other sources, so that we can fully initiate the model; third, an appropriate country and aggregation level need to be determined based on data availability and computation capacities; and finally, a full set of reliability weights in the objective function need to be selected in order to obtain meaningful solution from the model. We will discuss those issues one by one in five steps below.

3.1 Obtaining initial estimates for all bilateral trade variables in the model from observed or derived trade statistics

In east bound trade, initial estimates can be directly obtained from existing bilateral trade statistics for four sets of variables in the model. They are China's direct exports to partner countries ($DX0_{it}^{Ch,r}$), Hong Kong's total exports to partner countries ($TX0_{it}^{s,HK}$), and partner's total imports from China ($TM0_{it}^{s,CH}$) and imports of product originated from Hong Kong ($DM0_{it}^{s,HK}$). Similarly, there are also four sets of variables have initial estimates directly from existing data in westbound trade. They are partner countries total exports to Hong Kong and direct exports to China ($TX0_{it}^{s,HK}$ and $DX0_{it}^{s,CH}$), and China and Hong Kong's total imports from partner countries ($TM0_{it}^{s,CH}$ and $TM0_{it}^{s,HK}$). All China and Hong Kong reported trade statistics are obtained from China Custom authorities and Hong Kong Census and

Statistical Department in HS 8-digit details. All partner countries reported data are downloaded from UN COMTRADE at HS 6-digit level.

We also obtain initial estimates of Hong Kong's re-exports by origin and destination ($RX0_{it}^{sr}$) from Hong Kong re-exports statistics provided by Hong Kong Census and Statistical Department in HS 8-digit details. However, there are still nine sets of variables need initial estimates before the model can be implemented. There are four sets each for eastbound and westbound trade respectively, plus Hong Kong re-export markup ($RXM0_{it}^{sr}$). However, if we can obtain initial estimates for $RXM0_{it}^{sr}$ and also know fob/cif margin for all bilateral routes, then rest of the eight set variables all can be derived from existing trade statistics based on accounting identities specified in the optimization model.

The four sets unobservable variables in eastbound trade are China's total exports to partner countries ($TX0_{it}^{CH,r}$), Hong Kong's domestic exports to partner countries ($DX0_{it}^{HK,r}$),⁴ partner countries' direct imports from China ($DM0_{it}^{CH,r}$), and partner countries' total imports from Hong Kong ($TM0_{it}^{HK,r}$). Their initial estimates can be derived from observed data according to equation (2), (3), (4) and (5) respectively. The four sets unobservable variables in westbound trade are Hong Kong's imports from partner countries for domestic use ($DM0_{it}^{s,HK}$), China's direct imports from partner countries ($DM0_{it}^{s,CH}$), and partner countries' total exports to China and their exports for Hong Kong's domestic market ($TX0_{it}^{s,CH}$ and $DX0_{it}^{s,HK}$). Their initial estimates can be computed from observed data according to equation (7), (8), (9) and (10) respectively.

The initial estimates for bilateral trade variables between Hong Kong and China can be obtained from existing trade statistics reported by China and Hong Kong or calculated from observed trade data in the same fashion as unobserved variables in east and west bound trade according to equations (11) to (14). The observed statistics are $DX0_{it}^{CH,HK}$, $TX0_{it}^{HK,CH}$, $TM0_{it}^{CH,HK}$, and $DM0_{it}^{HK,CH}$. The only difference is that $TX0_{it}^{CH,HK}$ is China's actual exports to Hong Kong, equals its direct exports to Hong Kong minus all its re-export to other countries via Hong Kong.

In summary, there are eight set variables each in both eastbound and westbound as well as China and Hong Kong bilateral trade, four of them in each direction can be obtained directly from existing reported trade statistics. The remaining four sets unobservable variables we have four sets of equations in each trade direction, therefore, as long as we can obtain estimates for Hong Kong's re-exports markup ($RXM0_{it}^{sr}$) and fob/cif margins (cif_{it}^{sr}), all variables in the optimization model specified in this paper are fully initialized.

3.2 Calculate initial Hong Kong re-export markup rates

⁴ Although Hong Kong Census and Statistics Department also publishes Hong Kong's domestic exports to all its partner countries, but the definition is different with what we defined in this paper. We include Hong Kong's re-exports markup into Hong Kong's domestic exports.

The initial estimation of Hong Kong re-export markup rates follows the spirit of Feenstra *et al* (1998, 1999), the SAS programming procedures of which are documented in Chapter 2 of Yao (2000). While Feenstra *et al* (1998, 1999) only reports overall markup rates for China trade with the US and a few other selected countries, Yao (2000) is able to produce markup rates at 6-digit HS commodity and individual country levels. Yao (2000) also provides the markup rates tailored for trade data reconciliation in the GTAP version 5 database. This paper uses the same methodology and updated SAS procedures to estimate the average 2002-04 markup rates and their trade weighted standard deviations as the initial inputs for the mathematical programming model.

The key features of Feenstra *et al* (1998, 1999) include:

1. They use very detailed China and Hong Kong trade data at both the commodity level (SITC for early years and 6-digit HS for 1994 and onward) and country level. As a result, the markup rate estimates are also at the same detailed levels. The overall markup rate is just weighted average of those disaggregate markup rates.
2. The Hong Kong import data does not have information on the final destination countries but with China trade data, which identifies the final destination countries and origin countries that go through Hong Kong, they are able to produce better markup rate estimates for China-originated goods, but for China-bound goods, the markup rate estimates do not show any regular patterns.
3. The markup rate estimates are sensitive to outliers. By assuming that Hong Kong cannot re-export significantly more than its imports in the same year, records with re-export quantity more than double import quantity are treated as outliers and thus are deleted from the markup rate calculations.
4. Three methods produce three sets of markup rates and their aggregate values coincide with findings from JCCT (1995), which are based on the analysis of Hong Kong trade data only, Hong Kong Census surveys and Fung and Lau (1998) interviews. They reconcile all three sets of markup rates with precise economic interpretations. Specifically, **Method A** markup rates refer to those based on source generic Hong Kong import unit value but destination specific Hong Kong re-export unit value, and coincide with JCCT (1995) findings; **Method B** markup rates are based on Hong Kong import and re-export unit values both of which are source or destination generic, and coincide with Hong Kong Census survey results; and coinciding interview results reported in Fung and Lau (1998), **Method C** markup rates are based on Hong Kong import unit value (adjusted with China export data) and Hong Kong re-export unit values, both of which are source or destination specific and therefore are more accurate for China-US trade.

Markup rate is defined as the share of the value added by Hong Kong middleman in the total re-export value, or M_2 in Feenstra *et al* (1998). Let the unit-value of Hong Kong re-export be denoted by $PM_i = VM_i / QM_i$ where VM_i is the value and QM_i is the quantity of imports, and i denotes the HS codes. Let the unit-value of Hong Kong re-exports be denoted by $PX_i = VX_i / QX_i$, where VX_i is the value and QX_i is the quantity of re-exports. Thus the

relationship between the aggregate markup rate (RXMR) and disaggregate markup rate ($RXMR_i$) can be shown by the following formulas,

$$\begin{aligned}
 RXMR &= \frac{\sum_i (PX_i QX_i - PM_i QX_i)}{\sum_j PX_j QX_j} = \sum_i \left(1 - \frac{PM_i}{PX_i} \right) \left(\frac{PX_i QX_i}{\sum_j PX_j QX_j} \right) \\
 &= \sum_i RXMR_i \left(\frac{PX_i QX_i}{\sum_j PX_j QX_j} \right) = \sum_i RXMR_i \frac{RX_i}{\sum_j RX_j}
 \end{aligned} \tag{21}$$

The above formula shows that when using this definition, re-export values should be used as compatible weights.

For the purpose of using the programming model to solve for the final markup rate estimates, standard deviations are needed to measure the scope of variations of the estimates, and to inform the model how much adjustment should be allowed. The trade weighted variance and standard deviation of the markup rates are given as:

$$Var(RXMR_i) = \frac{\sum_k RX_k (RXMR_k - \overline{RXMR_i})^2}{\sum_j RX_j}, \text{ and } STD(RXMR_i) = \sqrt{Var(RXMR_i)} \tag{22}$$

where indexes j and k represents the group of 6 digit HS codes that in GTAP sector i , and were used to estimate the GTAP sector level mean markup rates, and again, the re-export values are chosen as weights to calculate the weighted variance.

To have better estimates for the trade weighted mean and variance of the markup rates, we first add up the annual data on Chinese exports, Hong Kong imports and re-exports over the years 2002, 2003 and 2004. So the markup rates should be interpreted as the trade weighted average over the three years. Both China and Hong Kong data are in 8-digit HS codes, but only comparable at 6-digit level. When calculating the Method A markup rates, only Hong Kong data is used and therefore, markup rates are at the 8-digit HS level. But in Method C markup rate estimation, we need to combine the Chinese export data with Hong Kong data. Because China and Hong Kong trade data are comparable only at the 6-digit HS level, Method C markup rates are estimated at 6-digit HS level. As final outputs, markup rates are aggregated into GTAP sector and trade region levels. To fully reflect the extent of markup rate spread over commodities, their variances and standard deviations are also calculated over 6-digit HS codes for a given pair of GTAP origin and destination countries at the GTAP sector level.

All initial markup rate estimates are Method A markup rates except for China originated goods, which have Method C markup rates. Method C could also apply to China bound goods when the unit values of Hong Kong re-exports to China are adjusted with Chinese import data, but we choose not to do so because Method A markup rates for China

bound goods do not show any regular patterns over years and it is not worth the efforts to improve it with Method C.

After obtaining those estimates at the level of GTAP sector and region, we replace negative markup rates with zero entries to keep them consistent with our mathematical programming model specifications, which do not permit negative values because of the counter intuitive nature of the negative markups at the aggregated level.⁵ However, eliminating negative values only slightly increases the overall markup rates for goods of China origin from 29.3% to 31.0%, and increases the overall markup rates for goods of China origin destined to the US from 32.3% to 32.8%. For goods destined to China, however, the increase due to removing the negative values are quite big, but they still lie within or close to the range of surveys by the Hong Kong Census and Statistical Department as reported in Table 2.6 of Fung *et al* (2006), or within the range of unreported initial estimates for the westbound US-China trade over 2001-05.

3.3 Bilateral trade cost and estimates of cif/fob margins

The procedure to estimate cif/fob margins for trade between China & Hong Kong and their trading partners is similar to what used for constructing the GTAP database which incorporates transport margins for all bilateral trade flows for each GTAP merchandise sector.

As discussed earlier, one source of discrepancies in reported trade flows is the recorded costs associated with shipping goods. These costs are generally recorded by the importing country while not included in the exporter's customs value at the port of origin. Although shipping costs alone are a minor contributor to the overall discrepancies found in bilateral trade statistics (Ferrantino and Wang, 2007), failing to take into account these cost in our model presents a problem for consistency and accuracy in the estimation of re-export markups. Bilateral transport margins can vary considerably by sector and trading partner. We therefore control for transportation cost on a bilateral basis when we initialize the model.

Attempts to impute transportation cost from the same trade data that suffers from reporting inaccuracies rarely yield credible estimates of shipping margins. We draw our transportation margin estimates directly from shipping cost information that is recorded and compiled consistently with the traded goods. The data we use is primarily from the U.S. Census Bureau on foreign trade statistics. According to U.S. Census' definition, the cif (cost, insurance, and freight) value represents the landed value of the merchandise at the first port of arrival. It is the sum of two components of the traded values: the "customs value" and the "import

⁵ The existence of negative markup rates at commodity level for a particular year does have its justification in the theories of intermediation, as discussed in section 2.1 of Feenstra and Hanson (2004). However, the same authors also attribute the negative markups at the aggregate level, say, at the 1-digit SITC level, to errors in markup rate calculations (Data Appendix C in Feenstra and Hanson, 2005) and do not accept those negative numbers in their econometric work. We share the same sentiments with them when replacing the negative aggregated markup rates with zeros in our mathematical programming model, though technically our model can handle the negative markup rates.

charges”. Consistency of transport cost is maintained when the cif value is computed by adding import charges to the customs value which excludes U.S. import duties. Import charges represent the aggregate cost of all freight, insurance, and other charges incurred in bringing the merchandise from alongside the carrier at the port of exportation in the country of origin and placing it alongside the carrier at its first port of entry. For overland shipments originating in Canada or Mexico, such costs include freight, insurance, and all other charges, costs and expenses incurred in bringing the merchandise from the point of origin where the merchandise begins its journey to the United States in Canada or Mexico to the first port of entry. The difference between the customs value and the cif value is expressed as a cif/fob ratio in the discussion that follows. Having the comprehensive commodity and partner coverage, this extensive dataset permits us to calculate GTAP sector aggregates from highly detailed bilateral commodity trade data.

The trade data with corresponding transport cost information is available at the most detailed level (10-digit HS) for all merchandise trade and for all U.S. trading partners. Thus we are able to calculate cif/fob margins for all U.S. trading partners directly. In the GTAP database (version 7) global merchandise imports valued at cif prices is 4 percent greater than merchandise exports valued at fob prices. This would imply that transport services add 4 percent to traded goods. This estimate reflects a trade-weighted average of all bilateral margins where some goods require a higher transport margin and others a lower margin depending on the characteristics of the product and the trading partner. Bulk goods with low unit values such as coal, iron ore, hides and skins, and bananas have higher transportation margins in the range of 20 to 40 percent. The cost of shipping raw or bulky-type goods is relatively expensive compared to goods with a high unit value which can be shipped in compact forms. Goods with high unit value such as computer components, precious metals, and jewelry commonly have transportation margins below 1 percent. However, within each aggregate GTAP sector there are both high-unit value goods and low unit value good which largely affects the range of bilateral aggregate margins. Thus longer distance between partner pairs does not necessarily correspond to a higher margin at the aggregate sector level.

The bilateral sector margins between China (or Hong Kong) with a particular partner are calculated using their bilateral trade as weights to sum up their corresponding transport margins estimated from US Census data set at the 6-digit HS level. Because of differences in commodity composition of trade flows, bilateral cif/fob margins for any aggregate sector will vary. For example, bilateral margins for the machinery and equipment sector (table 1) fall above or below the global merchandise average of 4 percent. The bilateral margins at the aggregate sector level are largely determined by the detailed content of the underlying bilateral trade flows. High unit value goods such as turbo-jets and other high-technology components (belonging to HS categories 8409-8411) can be shipped long distances even by air because shipping cost is relatively low. Timeliness of delivery is critical for such high value goods. The transportation margins (cif/fob) for this HS subgroup for all U.S. partners is 1.016. In contrast, another subgroup of machinery and equipment items such as air conditioners, pumps, fans with lower unit values have a higher cif/fob margin (1.041).

(Insert Table 1 here)

Each exporting country differs by its proportion of high value and low-value content it supplies which then has implications for the aggregate bilateral transport margin. To illustrate this how content differs we use the two HS categories shown in table 1 and show how the trade ratio of low value (LV) to high value (HV) goods differs substantially by exporting country. Generally the content of developing countries' manufactures differs from that of high income countries within any aggregated sector. For example Japan, Canada, Germany and the United Kingdom export a higher proportion high-value machinery and equipment than do China, Hong Kong, Mexico, and India. In fact China exports nearly 9 times more of the low-value category in machinery equipment than for the high-value category. Because of the higher transport margins on low-value goods, China's transport margin for exports is relatively high for its aggregate machinery and equipment sector (1.066). Although Brazil, India, and Mexico export a similar proportion of low value machinery and equipment, the aggregate cif/fob margin for Mexico is substantially lower (1.011) than for India and Brazil. This is largely because of the close proximity to the United States where efficient ground transportation is relatively cheap in comparison to ocean shipping by vessel transportation required for India and Brazil.

(Inset Table 2 Here)

Bilateral cif/fob margins for the other manufactures (table 2) differ substantially more than other aggregate GTAP sectors because of the wide range goods in this sector. This sector is comprised of goods with high unit value such as precious stones, metals, and jewelry and products with low unit value such as toys and sporting goods. As a result of mixing manufactures with precious minerals, the general pattern of higher transport margins for developing country goods does not hold. For example, India and South Africa have the lowest sector margins. The aggregate margins are driven primarily by the high proportion of precious stones such as diamonds and gold items in both countries exports of other manufacturers. Both countries have a very low LV/HV ratio as shown in table 2. China on the other hand exports a high proportion of low value goods (toys and sporting goods) and once again this is the reason for China's high transport margin (1.086) for the other manufactures sector as a whole.

The lowest and most uniform transportation sectors margins are those of the electronic equipment sector (table 3). Most countries supply a wide array of electronic items within the electronics sector where there is no clear specialization. These high-technology goods have some of the highest unit values of all merchandise goods. In addition, computer components such as chip sets and circuit boards are most often transported by air rather than vessel because of the time sensitive nature of these goods in the supply-chain management. Slight bilateral differences arise because only from subtle differences in the electronic content such as the lower value products of microphone, speakers, telephones and parts which have a higher transport margins than computer components. For example Costa Rica supplies a higher content of high-value computer chip sets than does India and China, making its aggregate margin for electronic equipment lower.

(Inset Table 3 Here)

A full set estimate on transport margins between China (Hong Kong) and their trading partners is required to initialize the model, including many non-U.S. bilateral trade flows. To

complete the estimation, specific margins are first calculated at the 6-digit HS level for all HS categories from U.S. Census' data but are grouped into two sets. One set is for countries that border each other and the other is for non-bordering countries. Then cif/fob margins for each route between China & Hong Kong and their trading partners at each GTAP sector is calculated as associated trade flow weighted average from 6-digit HS level based on whether the pair borders each other. For instance, the same 6-digit HS margin between U.S. and Mexico is applied to China and Hong Kong's trade, while the same 6-digit HS margin between U.S. and China is used to China and Brazil trade. Table 4 lists aggregate cif/fob margins for China's major exporting sectors to the U.S. and other its major partner countries. China's cif/fob margins with Hong Kong are considerably lower than trade with other partners due to the close proximity. There are some variations by importer due to the content of trade. We also assume that trans-Pacific and trans-Atlantic trade routes for the same goods would have the same margins. Although we do not have route specific information on freight rates, it is reasonable to assume that international shipping services are supplied by transportation firms outside the U.S. and that the same carrier shipping machinery from China to Brazil likely provides shipping services for goods shipped from China to the United States. Thus can we assume transports margins for the same goods would be similar as good carried on similar vessels from China to Brazil as China's exports to the United States.

(Inset Table 4 Here)

3.4 Determine appropriate country and commodity aggregation level based on the issue at hand and data availability

Because one of the objectives of this study is to produce Hong Kong re-exports adjusted trade flows contributing to version 7 GTAP database, therefore trade data reported by China, Hong Kong and their partners were aggregated from 8 and 6 digit HS to the 42 GTAP commodities respectively.

There are 215 countries identified in the GTAP global bilateral trade data base, while only 159 countries reported their exports to or import from China and Hong Kong during 2004.⁶ To determine the country aggregation used in our optimization model, we first aggregate all the non-reporting country into one block to be consistent with the model assumption that only one partner country do not report their trade with China and Hong Kong. Then use the difference between China reported imports (exports) and the sum of all partner reported exports (imports) adjusted by associate fob/cif margin to approximate the partner reported data for this aggregate non reporting country block. Then we use two cut off criteria to separate the 159 reporting country into two blocks. The first block has 95 countries except China and Hong Kong, including all single countries in version 6 GTAP database and the sum of exports from China and Hong Kong to the world greater than 300 million dollars in 2004 identified either by China and Hong Kong reported data or their partner reported data.

⁶ There are about 120 countries reported their trade with China and Hong Kong in 2004 in current WITS with missing data for China's several important trading partners such as Viet Nam. Therefore, additional data for 2002 to 2005 pulled directly from UN COMTRADE database were also used and growth rates between 2002 and 2003 were calculated at the 6-digit HS level to project missing data in 2004 before being aggregated into GTAP sectoral classifications.

The selected model country list and initial value of corresponding model variables for eastbound and westbound trade are listed Table 5 and Table 6 respectively. The second block is consisted of 64 remaining reporting countries. Their names are listed in Appendix Table A.

(Insert Table 5 and Table 6 here)

Although the initial estimates listed in Tables 5 and 6 still suffered from several unsolved data problems,⁷ they still show several interesting features of the data. First, reported westbound trade seems less problematic than reported eastbound trade, reflected by the more volatile statistical discrepancies in eastbound trade. The overall discrepancies are 5.8 percent in eastbound trade and 0.8 percent in westbound trade, there are 14 of the 98 reported bilateral routes in the model with more than 100 percent statistical discrepancies in the eastbound trade, while only one route in the westbound trade sees such large discrepancies. In the other hand, there are only nine bilateral routes in eastbound trade with less than five percent discrepancies, while more than 15 routes in the westbound trade have small discrepancies. Second, trade with developing country partners shows greater discrepancies than developed countries in general, reflecting the poor quality of data reported from those nations. Finally, extremely large discrepancies are usually come from partners only have small trade values with China and Hong Kong, such as Benin, Nigeria, Togo and Kyrgyz Republic in eastbound trade and Cambodia in westbound trade. The combined exports reported by China and Hong Kong are usually small in the 14 bilateral routes with more than 100 percent discrepancies in eastbound trade, and in all those countries their reported imports less significantly than what China and Hong Kong reported exports to them.

There are three types of balance of trade reported in Table 5. They are China and Hong Kong officially reported trade balance with their partner countries (difference between China and Hong Kong reported exports and imports before any adjustment), partner countries officially reported trade balance with China and Hong Kong (difference between partner reported exports to and imports from China and Hong Kong before any adjustment), and balance of trade after initial Hong Kong re-exports and fob/cif adjustments. While only adjusted trade balance are listed in Table 2 but calculated in an opposite direction, i.e. they should have a same absolute value with what reported in Table 5, but with an opposite sign. As expected, China's trading partners reported much larger trade deficits with China than China reported trade surpluses with its partners. The value reported by partner is \$327.9 billion, while the official statistics from China are only \$11.3 billion. More strikingly, if excluding Hong Kong, other partners reported a deficient with China at \$324.5 billion, while China also reported a trade deficient of \$7.9 billion with these partners! Most the initial adjusted trade balance fills between those two numbers. For example, the United States reported a \$174 billion trade deficit with China, while China only report about 80.4 billion trade surplus with the United States. This number after initial adjustment for Hong Kong re-exports markup earnings and

⁷ For example, the no reporting partners block has a very large discrepancy in east bound trade, because the difference between China and Hong reported total exports with the sum of all reporting partner's imports is too small relative to China and Hong Kong reported exports to these non reporting countries. This implies that even China and Hong Kong reported total exports and imports also subject to reporting errors which should be reconciled at global level with all importing and exporting country in the world.

fob/cif margins become 109.5 billion, 36 percent higher than the Chinese data, but 37 percent lower than data reported by the United States.

With all variables in the model has an initial value now, there is only one issue left before we can solve the optimization model: How the reliability weights in the objective function ($wdx_{it}^{sr}, wtx_{it}^{sr}, wdm_{it}^{sr}, wtm_{it}^{sr}, wrxm_{it}^{sr}$ in equation 18) should be determined, which will determine which and how much of the initial estimates should be adjusted and it is the topic in next section.

3.5 The choice and estimation of reliability weights

From statistical theory point of view, the best way to systematically assign reliability weights in the objective function is to obtain estimates of the variance-covariance matrix of the initial trade flow statistics. Then the inverted variance-covariance matrix may be justified as the best index of the reliability of entries in the trade flow matrix. However, the lack of consistent historical data often makes the estimation of the variance-covariance matrix associate with the initial trade flow statistics very difficult to implement. For example, the common practice in SAM balancing exercises is assign different degree of subjective reliabilities to the initial entries of the matrix follow the method proposed by Stone (1984),⁸ almost no attempt to date has been made to statistically estimate data reliability such as error variance of the initial estimates from historical data, except Weale (1989), who developed a statistical method that uses time series information on accounting discrepancies to infer data reliability in a system of national accounts. Theoretically speaking, a similar statistical method can be applied to the reporting discrepancies of bilateral trade data to derive those variances associated with international trade statistics.

Trade data reported by each country and its partners are often used in international economic literature to check the quality of trade statistics. An approximate match of mirror statistics implies trade data reported via that route are reliable. Therefore, an analysis of discrepancies between two "reported" trade data of the same trade flows may provide a means of determining data reliability and historical mirror trade statistics could be used as a major data source to estimate the variance of reported bilateral trade statistics.

Auto regression with dummy variables

Assuming the discrepancies in any pair of mirror trade statistics are a function of a systematic bias, last period's discrepancies and N dummy variables plus an error term as follows:

$$e_{it} = a_i e_{it-1} + b_i^0 + \sum_{k=1}^n b_i^k D^k + \mu_{it} \quad (25)$$

where e_{it} is the mirror trade statistics discrepancies at year t, b_i^0 is the symmetric bias, and μ_{it} is the random error term, D^k 's are dummy variables represent events have a significant impact on the reporting practice in the two data reporting countries such as change of commodity classifications, implementing better custom information systems or enforcing effective anti-

⁸ Stone proposed to estimate the variance of x_{ij}^0 as $\text{var}(x_{ij}^0) = (\theta_{ij} x_{ij}^0)^2$, where θ_{ij} is a subjective determined reliability rating, expressing the percentage ratio of the standard error to the initial estimates of x_{ij}^0 .

smuggling programs. The autocorrelation coefficient a_i and the variance $\text{Var}(\mu_{it})$ can be taken as indicators of magnitude of the measurement errors. The variance of initial trade statistics thus may be derived as follows: since e_{it-1} and μ_{it} are independent,

$$V(e_{it}) = V(a_i e_{it-1} + b_i^0 + \sum_{k=1}^n b_i^k D^k + \mu_{it}) = V(a_i e_{it-1}) + V(\mu_{it}) = a_i^2 V(e_{it-1}) + V(\mu) \quad (26)$$

At stationary assumption in long run, $V(e_{it}) = V(e_{it-1})$

$$V(e_{it}) - a_i^2 V(e_{it}) = V(\mu) \quad (27)$$

Therefore

$$V(e_{it}) = \frac{V(\mu)}{(1 - a_i^2)} \quad (28)$$

As long as we have enough historical mirror trade statistics and sufficient knowledge on the change in related country's trade reporting system to estimate $V(e_{it})$ for each pair of mirrored trade variables in our optimization model, then they can be assigned as weights in equation (18), the objective function. Although theoretically elegant and doable, however, the historical data and knowledge on the changes in related country's trade reporting system are too demand, make such method less attractable in large empirical applications like ours that involves many partner countries. Therefore, we adopted following two types of reliability indexes as the alternative.

Route Reliability indexes

As described earlier, in adjusting inconsistent bilateral trade flow statistics to satisfy the consistency requirements, it is crucial for the reconciliation procedure to more favorable towards changing the less reliable route than the more reliable route. For example, past statistical information suggested that the US-Japan trade is one of the most consistently reported trade flow. Thus, minor or no adjustment is needed on this particular route while more adjustment should occur where there is less certainty about the reported trade flow. Because a small discrepancy in mirror trade statistics may indicate a reliable trade route, while a large discrepancy may indicate unreliable reported data, mirror statistics and their discrepancies also directly provide useful information to construct some sort of reliability index to inform the model how the initial estimates should be adjust in the reconciliation process.

In fact, when we assign initial estimates for the 16 sets of trade flow variables in both east bound and westbound trade in the optimization model either directly from reported trade statistics or by derivation from them, we also obtain 8 sets of mirrored trade data. The discrepancies computed from each mirrored pair divided by corresponding sum of mirrored flows thus can be used to construct an index which reflects the reliability of the associate initial estimates of the reported

trade flows in some extent, although we are not sure how large the associated variance really may be. Using mathematical notations:

$$PDX_{it}^{cs} = PDM_{it}^{cs} = 2 \times ABS\left(\frac{cif_{it}^{cs} DM0_{it}^{cs} - DX0_{it}^{cs}}{DX0_{it}^{cs} + cif_{it}^{cs} DM0_{it}^{cs}}\right) \quad (29)$$

$$PTX_{it}^{cs} = PTM_{it}^{cs} = 2 \times ABS\left(\frac{cif_{it}^{cs} TM0_{it}^{cs} - TX0_{it}^{cs}}{TX0_{it}^{cs} + cif_{it}^{cs} TM0_{it}^{cs}}\right) \quad (30)$$

$$PDX_{it}^{sc} = PDM_{it}^{sc} = 2 \times ABS\left(\frac{cif_{it}^{sc} DM0_{it}^{sc} - DX0_{it}^{sc}}{DX0_{it}^{sc} + cif_{it}^{sc} DM0_{it}^{sc}}\right) \quad (31)$$

$$PTX_{it}^{sc} = PTM_{it}^{sc} = 2 \times ABS\left(\frac{cif_{it}^{sc} TM0_{it}^{sc} - TX0_{it}^{sc}}{TX0_{it}^{sc} + cif_{it}^{sc} TM0_{it}^{sc}}\right) \quad (32)$$

Where indexes “c” is indexed over set {CH, HK} and variable with a prefix “P” are reliability index for that variables.

All these reliability indexes defined above have a value between 0 and 2, defined in the similar spirit of Ferrantino and Wang (2007). A smaller value of the indexes indicates the initial estimates are relatively reliable for the associated trade route. The weights in the objective function (equation 18) of the model can be assigned by multiplying these indexes by their corresponding initial values, e.g., $wtx_{it}^{sr} = PTX_{it}^{sr} \times TX0_{it}^{sr}$. With such a weighting scheme, we also make the model to be biased towards changing initial estimates of those unreliable trade routes more than those reliable ones in the reconciliation process, because a larger index makes the weights larger thus adjustment of the corresponding initial estimates has a smaller contribution to the value of the objective function and will be adjusted more in the reconciliation process. For instance, China-Japan trade in both directions will adjust less proportionally than China-Togo trade, because China and Togo reported trade has a much larger absolute discrepancy than China and Japan reported trade.

Reporter reliability index

However, reliability weights defined above only consider the relative quality of initial estimates among all the bilateral routes, such weights treat the reported trade statistics from both reporters equally and do not distinguish which reporter is more reliable. In the case there is very unreliable reporter in the pair, it may adjust the reliable data reported by the partner too much thus loss original accurate information from the reliable partner. This is undesirable. To correct this problem, a reporter’s reliability index needs to be developed. Such an index should be able to dealing following three issues properly:

The first issue is related to the difference of reporting countries in their ability to report bilateral commodity trade. Variability in reporting quality across countries is highly relevant information for the problem we try to solve in our proposed modeling approach. As

discussed earlier, the adjustment process hinges heavily on relative reliability of the each reporting countries. An indicator of reporter reliability is basically a measure of how consistency a country reports its trade relative to their trading partners. However, judging a country's trade data based on a single bilateral flow alone is a poor reference, because a partner can misrepresent its trade thereby potentially discrediting a reliable reporter. Therefore, a good reporter reliability measure should take all reporting countries in the world into account in assessing a country's reporting reliability.

The second issue is what exactly should be captured by the reliability measure. The size of discrepancies could be incorporated into a measure of reliability such as relative route reliability index we defined earlier. However, placing emphasis on the magnitude of discrepancies only may over-penalize the reliability of a legitimate reporter. A poor reporter that makes an error for a given trade flow usually makes a similar error with other partners. For example a reporter that has mistaken the identity of one of its partners has implicitly made a mistake for others. It brings a systemic bias for that reporter. This type of problem should be detected and reflected in the reporter reliability measure without penalizing the reliable reporter.

The third issue is the capability of the measure to reflect both sector and country specific reliability information for each country as an exporter and as an importer. Countries typically have commodity specific strength and weaknesses. For example one exporting country may have an excellent reporting record on steel but at the same time is highly inconsistent in its reporting practice in organic chemical trade.

All three issues discussed above are effectively dealt with in the type of reliability index developed by Gehlhar (1996) where reporter reliability indices were used to make a discreet choice whether to disregard or accept reported trade flows. The index is calculated as the share of accurately reported transactions of a reporter's total trade using a threshold level. It assesses reporter reliability from a complete set of global reporting partners, captures the reporter's ability to accurately report without interferences from gross discrepancies in reporting, and contains exporter and importer-sector specific reliability information. Specifically, the importer-sector specific and exporter-sector specific reliability indexes are defined as:

$$RIM_{it}^r = \frac{100 \times MA_{it}^r}{\sum_s M_{it}^{sr}} \quad \text{where} \quad MA_{it}^r = \sum_{s \in AL_{it}^{sr} \leq 0.20} M_{it}^{sr} \quad AL_{it}^{sr} = \frac{|M_{it}^{sr} - X_{it}^{sr}|}{M_{it}^{sr}} \quad (33)$$

$$RIX_{it}^r = \frac{100 \times XA_{it}^r}{\sum_s X_{it}^{rs}} \quad \text{where} \quad XA_{it}^r = \sum_{s \in AL_{it}^{rs} \leq 0.20} X_{it}^{rs} \quad AL_{it}^{rs} = \frac{|M_{it}^{rs} - X_{it}^{rs}|}{M_{it}^{rs}} \quad (34)$$

where M_{it}^{sr} and X_{it}^{sr} are sector i imports and exports reported by country r and s in year t respectively, both measured at fob prices. Under such defined reporter reliability indexes, the size of the discrepancies becomes immaterial because inaccurate transactions are treated the

same regardless of the magnitude of the inaccuracy. The indexes have the flexibility of being implemented at the detailed 6-digit HS level and can be aggregated to any sector level. We computed such reporter reliability measures for China & Hong Kong and all their partners at the GTAP sector level. Major data are from UN COMTRADE with supplements from country sources.

After RIM and RIX calculated for each of the 99 countries including China and Hong Kong in the model for each GTAP sectors, the weights in the objective function (equation 18) of the model can be assigned by multiplying one minus these indexes by their corresponding initial values for each variable in the model. The complete set of weights in equation 18 is defined as follows:

$$wtx_{it}^{sr} = \lambda(1 - RIX_{it}^s)PTX_{it}^{sr} \times TX0_{it}^{sr} \quad (35)$$

$$wtm_{it}^{sr} = \lambda(1 - RIM_{it}^r)PTX_{it}^{sr} \times TM0_{it}^{sr} \quad (36)$$

$$wdx_{it}^{sr} = \lambda(1 - RIX_{it}^s)PDX_{it}^{sr} \times DX0_{it}^{sr} \quad (37)$$

$$wdm_{it}^{sr} = \lambda(1 - RIM_{it}^r)PDX_{it}^{sr} \times DM0_{it}^{sr} \quad (38)$$

$$wrxm_{it}^{sr} = \frac{STD(RXMR_{it}^{sr})}{RXMR_{it}^{sr}} RXM0_{it}^{sr} \quad (39)$$

Where λ is a scale parameter, $STD(RXMR_{it}^{sr})$ is defined by equation (22).

With such a weighting scheme, we also make the model to be biased towards changing those unreliable initial data more than those reliable ones in the reconciliation process. It means the reconciled solution from the model not only adjust less to the reliable routes than the unreliable ones, but also adjust more to the relative unreliable reporter than the relative reliable reporter in each trade route, although in a rough manner.

IV. Preliminary Results from the Model

The optimization model is coded in GAMS (Brooke *et al*, 2005), with more than 2.5 million equations and variables in its current aggregation. It was solved using barrier method of the Cplex solver (GAMS Development Corporation, 2005) in a 32 bit dell computer with 3 GB memory. There are 13 input data files, all automatically produced by three SAS programs.

Adjusted estimates for the sum of all sectors aggregated into 24 regions are listed in Tables 7 and 8 for eastbound and westbound trade respectively. To facilitate better comparison of trade flows before and after model adjustment, the corresponding aggregated initial data also presented in the same tables. Results for the 99-country details are listed in appendix tables B

and C. More detailed initial and model adjusted trade flows by countries and GTAP sectors are available from the author upon request.

(Insert Table 7 and Table 8 here)

4.1 Adjusted trade flow and balance of trade between China and its major trading partners

Table 7 reports model adjusted aggregate bilateral trade flow and balance of trade between China, Hong Kong and their major trading partners along with official trade balance reported by both sides. For eastbound trade, Chinese exports to the US have an upward adjustment of 15.7%; for Japan, it is 4.8%; for Korea, it is -5.5%; for EU 15, it is 7.7%; and for ASEAN, it is -13.3%. In terms of percentage change, the adjustments of exports to major destinations are well below those to Russia and Rest of Africa, etc. This indicates that though there is still room for the model to further adjust the data of Chinese exports to major partners, relative speaking, the quality of initial estimates is quite good as long as institutional factors that could distort official trade data are considered in initial data adjustments.

For westbound trade, the percentage adjustments are much smaller for the same major trading partners. For the US exports to China, the model adjustment is only 2.2%; for Japan, 0.3%; for Korea, 1.6%, EU 15, -0.9%; and for ASEAN, -0.7%.

Though these model adjustments of Chinese bilateral trade flow with major trading partners are modest or small, they may have big impacts on the model adjusted trade balance. For example, the model adjustment of China-ASEAN balance of trade is 744.8%, for China-Japan trade balance, it is 46.3%; for China-EU 15 trade balance, it is 17.3%; and for the all-too important China-US trade balance, there is additional 20.6% increase compared to traditional reconciliation, i.e., the initial estimates. In short, the model does make a difference, sometimes a big difference in reconciling trade flows and trade balances between China and its major trading partners.

Despite the percentage adjustment of balance of trade seems large for some of the bilateral routes, most of the adjusted bilateral balance of trade consistently lies between China and its partner officially reported data. For example, the model adjusted trade surplus for China is 127.3 billion dollars, which is significantly higher than China officially reported surplus of 11.3 billion⁹, but also significantly smaller than the 327.9 billion partners reported trade deficit with China. At bilateral level, for instance, the model adjusted trade balance between China and Canada is 6.8 billion dollars in China's favor, which lies between the 0.8 billion China reported trade surplus with Canada and 13.5 billion Canada reported trade deficit with China. Similarly, the model adjusted trade balance between China and the 15 member of European Union is 65.3 billion dollars in China's favor, which also lies between the 31 billion China reported trade surplus with EU 15 and EU 15 reported 98.8 billion trade deficits

⁹ The balance of trade data reported here are calculated from current model data base, which is different from the officially reported data because our model database excludes utility trade (such as electricity) and HS Chapter 98 and 99. There are also 36.9 billion Hong Kong re-exports of China originated products back to China did not count as China's imports as described in the text. Therefore, China's trade surplus in the model is lower than 32 billion, the official 2004 number reported by China.

with China (bottom section of table 7).

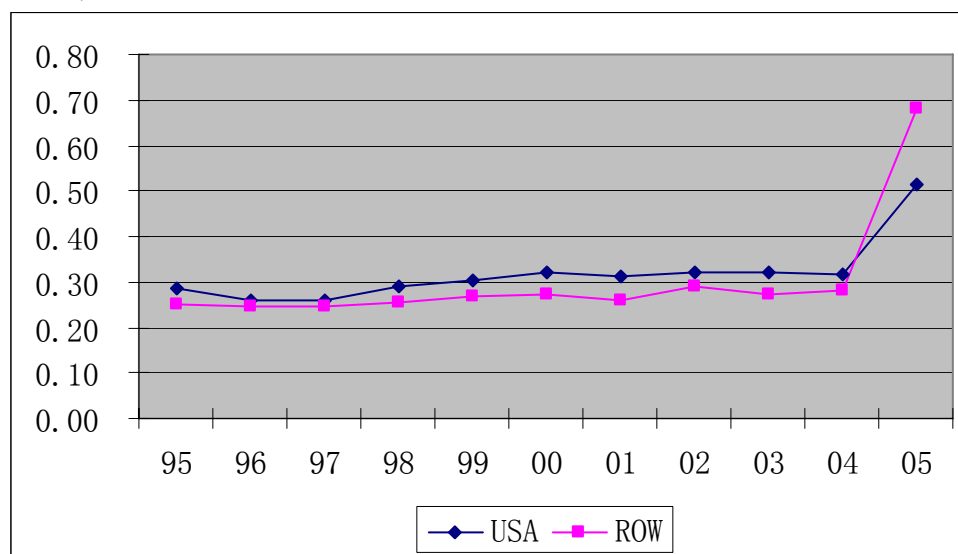
4.2 Adjusted Hong Kong re-export markup rates

The model does make a difference in adjusting the Hong Kong re-export markup rates. For eastbound re-exports, the differences are bigger, while for westbound re-exports, the differences are smaller. Specifically, the model increases the markup rate for Chinese goods re-exported to the rest of the world from 31.0% to 37.8%, while for goods from rest of the world re-exported to China, the markup rate is increased slightly from 7.9% to 8.9%. Because of some data and logical issues still unsolved in the model, which we will discuss later, the rightfulness of these adjustments is subject to further investigation. Nevertheless, the adjustments seem to be acceptable.

In terms of country breakdown, the model significantly increases the markup rates for all destination countries in eastbound trade. Among them, the China-US markup rate is raised from 32.8% to 37.6%. In comparison, for westbound trade, most countries, -- except Japan, Korea and ASEAN (China's top deficit countries/region) as well as Mexico, -- experience small increases in the markup rates for their goods shipped to China via Hong Kong.

To better understand our model results for eastbound trade, which experiences significant adjustment, we put the markup rates in perspective. Using the approach described in Section 3.2, we also calculate the markup rates for the past ten years (1995-2005), and as shown in Figure 1, a pattern has been revealed: China-US markup rates are consistently higher than the China-world markup rates and both are increasing over 1995-2004. In 2005, however, there is a sharp rise on both markups, and the China-world markup (68.2%) is higher than the China-US markup (51.6%). We still do not know what causes this sharp jump and the reverse of the relative size of the two markup rates, but they seem to lend some justification for our model results.

Figure 1 Hong Kong markup rates for re-export Chinese goods to the US and rest of the world, 1995 to 2005



The higher adjusted markup rates for Chinese goods may also have something to do with our model's treatment of the Hong Kong re-exports of Chinese goods back to China, totaling \$34.8 billion in Hong Kong trade statistics (or \$36.9 billion in Chinese Customs statistics).¹⁰ They are simply eliminated from the statistics of Hong Kong's re-exports, total exports and imports, but no similar adjustment has been made to China's direct exports to Hong Kong, because there are no such information available in Chinese official export data. As a result, adjustments have to be made to account for the deleted round-tripping trade flow, which in part lead to the upward adjustment of the re-export markup rates for the Chinese goods.

In terms of sectoral breakdown, in eastbound trade, most positive adjustments happen to only a few sectors, such as animal, food, beverage and tobacco products, as well as transportation and electronic products. In westbound trade, the biggest rise in markup rates go to animal products, wool silk and worm cocoons, and coal.

(Inset table 9 here)

Table 9 presents the initial and model adjusted Hong Kong re-exports as percentage of China's total exports. For eastbound trade, the model reduces overall share of re-exports via Hong Kong in total Chinese exports by 1.5% (from 13.2% to 11.8%). The sectors that are mostly affected are GTAP sector 42 (manufactures nec, -8.0%), sector 31 (paper products publishing, -6.2%), sector 12 (wool silk-worm cocoons, -5.3%), and sector 40 (electronic equipment, -3.7%). Similarly, for westbound trade, the overall share of Chinese imports via Hong Kong in total Chinese imports declines by only 1.1% (from 16.3% to 15.3%), but the impacts are concentrated mostly in two sectors: sector 28 (wearing apparel, -31.5%) and sector 26 (beverages and tobacco products, 11.1%).

4.3 Hong Kong re-exports earnings and retained imports

The first panel of table 10 summaries Hong Kong's earnings from its re-export China originated goods to other countries, from re-exports other countries' products to China, and from re-exports of commodities among other countries via Hong Kong by GTAP sectors. It shows that for all sectors combined, re-export earnings from Chinese goods are highest in value terms and also have the most adjustment in terms of the percentage change (21.9%), followed by earnings for re-exports of China-bound goods in terms of both value and percentage change (12.9%). For all other goods, their earnings and adjustment are the smallest in terms of value and percentage change (0.4%). Despite the difference in magnitudes, earnings in all the three categories have positive adjustments. Similar to discussions in section 4.1 on the round-tripping re-exports of the Chinese goods, the same explanation may also apply to the higher re-export earnings from

¹⁰ This may be quite true in real world trade. For example, shipments of forest products from northwest port of Dalian to Hong Kong by sea first, then transport to factories use these products in Shenzhen by truck may be a lot cheaper than direct transport the products from inland China to Shenzhen. However, the data show that the majority of these round tripping commodities are electronic equipment (17.8 billion), Other machineries (7.3 billion) and textiles (5.0 billion), there must be some incentive reasons to encourage exporters to do so.

Chinese goods.

Across sectors, the percentage changes in re-export earnings from China-bound goods vary the most, followed by earnings from Chinese goods. Adjustments in earnings from all other goods have the minimal variations across sectors in percentage terms.

Nevertheless, both the initial and the adjust estimates show that Hong Kong's re-export activities and their associated earnings are mainly concentrated on few differentiated manufacturing products. In eastbound trade, these products are: (1) electronic equipment, (2) other machinery and equipment, (3) other manufactures, (4) wearing apparel, (5) leather and sporting goods, (6) textiles, and (7) chemical, rubber and plastic products. These seven GTAP sectors account for 93 percent Hong Kong's markup earnings from re-exporting China originated goods to the world in the initial estimates, and 94 percent in the model adjust estimates. While electronics equipment and other machinery are the two major products that Hong Kong re-exports for other countries to China. Earnings from these two GTAP sectors constitute about 70 percent Hong Kong's markup earnings in westbound trade for both the initial and adjust estimates. Qualities of these products are usually more difficult to observe and more likely to require the service of intermediation to resolve information problems in trade (Feenstra and Hanson, 2004). Therefore, these estimates make good economic sense.

(Inset table 10 here)

The second panel of table 10 lists initial and adjusted estimates of Hong Kong's retained imports from all its trading partners excluding and including China by GTAP sectors. The initial estimates seem very close to the estimates for 2004 published by Hong Kong Census and Statistics Department at the aggregate level when excluding imports from China (68.7 and 72.9 billion U.S dollars respectively), while the model adjusted estimates are significantly larger. However, carefully comparing the initial and adjusted estimates, we find our current treatment of Hong Kong re-export China originated products to China in the model is an major contributing factor to such a results. Recall discussions on our model's treatment of the \$34.8 billion round-tripping Chinese re-exports. It is very possible that the exporters misreported to Chinese Customs that such exports are bound for some other final destinations via Hong Kong for incentive reasons, such as export rebates, but in fact these exports went back to China eventually as shown in both Hong Kong's re-exports and China's official imports statistics. Therefore, the model tends to over estimate Hong Kong retained imports and introduces bias to its estimates of Hong Kong re-exports markup rates. This is clearly shown by comparing the last column of table 10, Hong Kong re-exports China originated goods back to China, with initial and adjusted estimates of Hong Kong's retained imports. There are three sectors in the model adjusted estimates are significantly higher than these initial estimates. They are electronic equipment (up from 27.2 to 42.9 billion), chemical rubber plastic products (from 6.3.to 9 billion), and other machinery and equipments (up from 7.2 to 19.5 billion), while the corresponding Hong Kong re-exports from China back to China are 17.8, 1.5 and 7.3 billion respectively. Obviously, treat such round trip trade flows properly in the model will improve the accuracy of the final estimates.

4.4 Adjusted China's balance of trade at sector level

The first panel of table 11 presents initial and model adjusted net exports of China with all its trading partners with and without Hong Kong by GTAP sectors. There are several interesting features of the model adjusted estimates of China's net exports to the world. First, there is no sign change among China officially reported net exports, the initial, and model adjusted estimates, in both including and excluding Hong Kong, for all but one GTAP sectors (beverages and tobacco products). Further more, for the sector with sign changes, the sign of model adjusted net exports are the same with the officially reported net exports and consistent with people's intuition of China's comparative advantages (net exporter of beverages and tobacco products, which is labor intensive). Finally, by adjusting Hong Kong's re-exports back to China's total export and imports, the adjusted net trade flows show China's current comparative advantages in the world market more clearly. For instance, the adjusted net exports are significantly larger than China officially reported in most labor intensive products such as leather and sporting goods, wood products, other manufactures and certain technology-capital intensive goods such as electronic equipments. All these imply that Hong Kong's re-export activities facilitate China to fully realize its comparative advantages and the model did a reasonable job in adjusting China's net trade flows.

(Inset table 11 here)

China's trade balance with the United States by GTAP sector is presented in the second panel of table 11 as an example to illustrate the features of model adjusted bilateral net trade flows at sector level. It also shows that most model adjusted sector net trade flows lie between China and the U.S. officially reported statistics except few sectors, which are associated with either very small trade balance or China and US both reported surplus or deficit with each other. It is also interesting to note there are four GTAP sectors that the initial estimates on sector balance of trade were adjusted out of the range that reported by the two trading partners, while the model is able to drive the final estimates back (or closer) to its reasonable band (vegetable and fruits, other animal products, forestry, fishing), thus demonstrate the desirable property of the model in trade statistics reconciliation.

V. Concluding Remarks

This study constructed a mathematical programming model to estimate re-export markup and reconcile detailed bilateral trade statistics from China, Hong Kong and their trading partners. Five key steps to link the model with actual trade statistics are described. The model was applied to 2004 bilateral world trade data in GTAP sectoral classification to produce Hong Kong re-exports adjusted trade flows contributing to version 7 GTAP database. Preliminary result shows that the model is able to eliminate the statistical discrepancy efficiently and at the same time provides a positive re-exports markup estimates in both directions for all covered commodities. Hong Kong's re-export mark-up, each trading partner's exports and imports via Hong Kong as percent of the country's total exports to and import from China, and adjusted bilateral balance of trade among China, Hong Kong and their partner countries by each covered commodity are all part of the model solution.

In conclusion, the model provides a flexible tool to reconcile trade statistics from China, Hong Kong and their trading partners simultaneously. Advantages of the model are its flexibility in data requirement and its desirable theoretical and empirical properties. It can be applied to reconcile direct and indirect trade for other regions of the world where re-export activities create major discrepancies. It not only provides a tool for the preparation of global trade data in future versions of GTAP database, but also contributes to the methodological development to estimate and reconcile discrepancies in international trade statistics when re-export activities diminish the ability of a country to identify its partner countries correctly.

However, there are several caveats that need mentioning. First, we keep re-export statistics reported by Hong Kong Census and Statistics Department as constant in the model during the adjustment process, because it is the most reliable source to provide both origins and destinations of re-exports through Hong Kong. In reality, such statistics also subject to errors as other reported trade statistics.

Second, the model assumes both China and Hong Kong correctly report their total exports to and imports from the world, therefore, these totals enter the model as controlled constants. However, in the real world, the sum of partner countries reported trade with China and Hong Kong in some sectors may well exceed what China and Hong Kong reported as illustrated by the huge negative discrepancies in derived trade statistics for the no reporting country block in our model. Therefore, there is an inconsistency in the global level which can not be eliminated by the current model alone. To solve this issue, a global commodity equilibrium adjustment model is needed. It treats each country as both supplies and demanders for each commodity and reconciles each countries' total exports and imports statistics using equation (17) as its constraint to solve a set of global consistent total exports and imports (no bilateral trade data needed) for each commodity and every country, which then can be used as input to our current model to solve the bilateral details.

Third, we made our estimates on bilateral transport margins primarily from trade related shipping cost information from the United States and these estimates enter the model as constant parameters. The associated errors with these parameters may transmit through the model thus impact the accuracy of the re-export markup and bilateral trade flow estimates. Therefore, the numeric estimates reported in the paper should be interpreted with caution and sensitivity analysis should be conducted in future studies to check how changes in these fixed parameters of the model may impact results from its solutions.

Finally, current model only reconciles one year's bilateral trade data, to be consistent with the 2004 base year of version 7 GTAP database. However, a three year average may be more desirable. It not only smoothes any unusual annual variation of the bilateral trade data, reduce time difference of recording which might cause discrepancies, but also provides more non-zero entries in the trade flow matrix. This certainly will have a positive impact on the development of CGE-based trade policy analysis using future versions of GTAP database.

Table 1 Bilateral transport margins on selected U.S. import flows for machinery and equipment sector

	<i>High unit value (HV)</i>	<i>Low unit value (LV)</i>	<i>Traded content</i>	<i>Aggregate</i>
HS category	8409-8411	8413-8415	Ratio	Machinery&Equip
	cif/fob	cif/fob	LV / HV	cif/fob
Canada	1.010	1.013	0.42	1.012
Mexico	1.007	1.005	1.36	1.011
Costa Rica	1.014	1.128	0.12	1.025
Brazil	1.077	1.064	1.00	1.045
United Kingdom	1.005	1.034	0.20	1.031
Germany	1.018	1.029	0.74	1.034
India	1.070	1.087	0.92	1.052
China	1.036	1.088	8.91	1.066
Hong Kong	1.017	1.088	2.93	1.052
South Korea	1.026	1.062	2.01	1.055
Taiwan	1.034	1.064	2.94	1.046
Japan	1.028	1.030	0.68	1.033
Australia	1.020	1.050	0.71	1.034
South Africa	1.060	1.116	0.29	1.040
World	1.016	1.041	0.79	

Table 2 Bilateral transport margins on selected U.S. import flows for other manufacturers

	<i>High unit value (HV)</i>	<i>Lower unit value (LV)</i>	<i>Traded content</i>	<i>Aggregate</i>
HS category	7101-7118	9501-9508	Ratio	Other manufacturers
	cif/fob	cif/fob	LV / HV	cif/fob
Canada	1.006	1.056	1.27	1.034
Mexico	1.003	1.063	1.67	1.051
Costa Rica	1.002	1.000	0.67	1.043
Brazil	1.014	1.176	0.17	1.038
United Kingdom	1.004	1.044	0.50	1.014
Germany	1.008	1.045	1.43	1.034
India	1.004	1.093	0.01	1.005
China	1.039	1.088	9.63	1.086
Hong Kong	1.008	1.101	0.33	1.036
South Korea	1.027	1.037	1.37	1.049
Taiwan	1.052	1.073	17.37	1.093
Japan	1.005	1.020	5.58	1.039
Australia	1.002	1.029	0.63	1.022
South Africa	1.000	1.078	0.01	1.005
World	1.006	1.079	0.89	

Source: U.S. Census, foreign trade statistics using transport costs (c.i.f. / customs value).

Note: HS categories of high unit value goods consist of precious stones, metals, and jewelry categories in low unit value goods are primarily toys, sporting goods, and accessories.

Table 3 Transport margins on selected U.S. import flows for electronic equipment sector

<i>HS</i>	<i>8471</i>	<i>8518</i>	<i>8517</i>	<i>Aggregate</i>
	Computer/parts	Microphones/parts	Telephone/parts	Electronic equip.
Canada	1.006	1.004	1.005	1.004
Mexico	1.002	1.003	1.002	1.003
Costa Rica	0.996	1.020	1.017	1.012
Brazil	1.026	1.077	1.022	1.019
United Kingdom	1.017	1.022	1.016	1.022
Germany	1.020	1.026	1.011	1.020
India	1.039	1.093	1.057	1.026
China	1.024	1.067	1.022	1.023
Hong Kong	1.029	1.078	1.018	1.022
South Korea	1.025	1.043	1.028	1.016
Taiwan	1.022	1.053	1.028	1.019
Japan	1.018	1.034	1.018	1.022
Australia	1.018	1.018	1.015	1.020
South Africa	1.023	1.029	1.015	1.031
World	1.019	1.047	1.013	

Table 4 Bilateral aggregate transport margins for China's exports to its major trading partners

	<i>Footwear</i>	<i>Chemicals, Rubber, plastics</i>	<i>Electronic equipment</i>	<i>Machinery & equip.</i>
Canada	1.076	1.085	1.024	1.065
Mexico	1.109	1.084	1.025	1.048
United States	1.072	1.093	1.023	1.066
Costa Rica	1.080	1.080	1.028	1.067
Brazil	1.091	1.062	1.028	1.041
United Kingdom	1.091	1.095	1.026	1.066
Germany	1.094	1.082	1.022	1.054
India	1.080	1.054	1.020	1.046
Hong Kong	1.020	1.019	1.007	1.016
Taiwan	1.074	1.080	1.021	1.040
Japan	1.084	1.089	1.023	1.048
Korea	1.072	1.092	1.024	1.039
Australia	1.079	1.091	1.024	1.064
South Africa	1.075	1.084	1.029	1.059

Source: U.S. Census, foreign trade statistics using transport costs (c.i.f. / customs value).

References

Bohatyretz, Sandra and Bruna Santarossa, 2005, "Merchandise Trade Reconciliation Study: Canada-China, 2002 and 2003," *Canadian Trade Review*, No 3, International Trade Division, Statistics Canada, Ottawa, <http://www.statcan.ca/english/research/65-507-MIE/65-507-MIE2005003.htm>

Brooke, Kendrick, Meeraus, and Raman, 2005, "GAMS -- User's Guide" GAMS Development Cooperation, Washington, DC.

Byron, Ray P. 1978. "The Estimation of Large Social Account Matrix," *Journal of Royal Statistical Society*, A, 141 (Part 3), 359-367.

Canning, Patrick and Zhi Wang, 2005, "A Flexible Mathematical Programming Model to Estimate Interregional Input-Output Accounts." *Journal of Regional Sciences* 45(3):539-563, August.

Deardorff, Alan V. 1998, "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" In J. A. Frankel, ed., *The Regionalization of the World Economy*, pp. 7-22. Chicago: University of Chicago Press.

Jonathan, Eaton and Samuel Kortum, 2002, "Technology, Geography, and Trade," *Econometrica*, Vol. 70: 1741-1779, September.

Feenstra, Robert, Wen Hai, Wing Woo and Shunli Yao, 1998, "The U.S.-China Bilateral Trade Balance: Its Size and Determinants," *NBER Working Paper #6598*, June, National Bureau of Economic Research (NBER), Cambridge, MA

Feenstra, Robert, Wen Hai, Wing Woo and Shunli Yao, 1999, "Discrepancies in International Data: An Application to China-Hong Kong Entrepôt Trade," *American Economic Review*, vol. 89, no. 2, 338-343, May.

Feenstra, Robert C., James R. Markusen and Andrew K. Rose, 2001, "Using the Gravity Equation to Differentiate Among Alternative Theories of Trade," *Canadian Journal of Economics*, Vol. 34: 430-447, May.

Feenstra and Robert C. and Gordon H. Hanson, 2004, "Intermediaries in Entrepôt Trade: Hong Kong Re-Exports of Chinese Goods," *Journal of Economics and Management Strategy*, Vol. 13, No. 1, pp. 3-35.

Feenstra, Robert and Gordon Hanson, 2005, "Ownership and Control in Outsourcing to China: Estimating the Property-Rights Theory of the Firm," *Quarterly Journal of Economics*, p729-761, May

Ferrantino Michael J and Zhi Wang, 2007, "Accounting for Discrepancies in Bilateral Trade: The Case of China, Hong Kong, and the United States" ITC staff working paper, 07-04-A.

Fung, K C, 1996, "Accounting for Chinese Trade: Some National and Regional Considerations," *NBER Working Papers* 5595, National Bureau of Economic Research (NBER), Cambridge, MA

Fung, K C and Lawrence Lau, 1998, "The China-United States Bilateral Trade Balances: How Big Is It Really?" *Pacific Economic Review*, No. 3, October, pp. 33-47

Fung, K C and Lawrence Lau, 2001, "New Estimates of U.S.-China Bilateral Trade Balances," *Journal of the Japanese and International Economics*, Vol. 15, pp. 102-130

Fung, K C and Lawrence Lau, 2003, "Adjusted Estimates of United States-China Bilateral Trade Balances: 1995-2002," *Asian Economic Journal*, Vol. 14, May/June, pp. 489-496.

Fung, K C, Lawrence Lau, and Yanyan Xiong, 2006 "Adjusted Estimates of United States-China Bilateral Trade Balances—An Update," *Pacific Economic Review*, vol 11(3), pages 299-314, October

Gehlhar, Mark, 1996, "Reconciling Bilateral Trade Data for Use in GTAP" GTAP Technical Paper No. 10, Purdue University.

Harrigan, J. Frank 1990, "The Reconciliation of Inconsistent Economic Data: the Information Gain," *Economic System Research*, Vol.2, No.1, pp. 17-25

Harrigan, J. Frank and Iain Buchanan. 1984. "A Quadratic Programming Approach to Input-Output Estimation and Simulation," *Journal of Regional Science*, 24(3), 339-358.

Hong Kong Census and Statistical Department (HKCSD), *Hong Kong Monthly Digest of Statistics*, various issues, Hong Kong

Joint Commission on Commerce and Trade (JCCT), 1995, "Report of the 'Trade Statistics Subgroup'," Trade and Investment Working Group, Washington DC, October 17, <http://www.census.gov/foreign-trade/reconcile/china.html>

Ploeg, van der F, 1984, "General Least Squares Methods for Balancing Large Systems and Tables of National Accounts," *Review of Public Data Use*, 12, 17-33

Schindler, W. John and Dustin Beckett, 2005, "Adjusting Chinese Bilateral Trade Data: How Big China's Trade Surplus?" *International Journal of Applied Economics*, 2(2):27-55, September.

Stone, Richard, 1984, "Balancing the National Accounts: The Adjustment of Initial Estimates: a Neglected Stage in Measurement," in A. Ingham and A.M. Ulph (eds.), *Demand, Equilibrium and Trade*. London: Macmillan

Weale, Martin R, 1985, "Testing Linear Hypotheses on National Account Data," *Review of Economics and Statistics*, 67, 685-689.

Weale, Martin R, 1989, "Asymptotic Maximum-Likelihood Estimation of National Income and Expenditure," Cambridge, mimeo.

Yao, Shunli, 2000, *Three Essays on China's Foreign Trade*, unpublished PhD dissertation, University of California, Davis.

Appendix Table A Countries in the other reporting country block of the model

<i>Country number</i>	<i>ISO3</i>	<i>Country name</i>	<i>Country number</i>	<i>ISO3</i>	<i>Country name</i>
1	ABW	Aruba	33	KNA	St. Kitts and Nevis
2	AND	Andorra	34	LBY	Libya
3	ARM	Armenia	35	LCA	St. Lucia
4	AZE	Azerbaijan	36	LSO	Lesotho
5	BDI	Burundi	37	MDA	Moldova
6	BFA	Burkina Faso	38	MDV	Maldives
7	BHR	Bahrain	39	MKD	Macedonia, FYR
8	BIH	Bosnia and Herzegovina	40	MLI	Mali
9	BLR	Belarus	41	MNG	Mongolia
10	BLZ	Belize	42	MRT	Mauritania
11	BOL	Bolivia	43	MSR	Montserrat
12	BRB	Barbados	44	MUS	Mauritius
13	BRN	Brunei	45	NAM	Namibia
14	CAF	Central African Republic	46	NCL	New Caledonia
15	CIV	Cote d'Ivoire	47	NER	Niger
16	CMR	Cameroon	48	NIC	Nicaragua
17	COK	Cook Islands	49	NPL	Nepal
18	CPV	Cape Verde	50	OMN	Oman
19	DMA	Dominica	51	PNG	Papua New Guinea
20	ERI	Eritrea	52	PYF	French Polynesia
21	ETH	Ethiopia(excludes Eritrea)	53	QAT	Qatar
22	FJI	Fiji	54	RWA	Rwanda
23	GAB	Gabon	55	SEN	Senegal
24	GEO	Georgia	56	SLE	Sierra Leone
25	GIN	Guinea	57	SLV	El Salvador
26	GMB	Gambia, The	58	STP	Sao Tome and Principe
27	GRD	Grenada	59	SUR	Suriname
28	GRL	Greenland	60	SWZ	Swaziland
29	GUY	Guyana	61	SYC	Seychelles
30	HND	Honduras	62	TTO	Trinidad and Tobago
31	ISL	Iceland	63	VCT	St. Vincent and the Grenadines
32	JAM	Jamaica	64	WSM	Samoa