# Voorburg Conference Notes

# Day 1 – September 13, 2022

**Opening Remarks & Meeting Agenda Overview**  
**Voorburg Group co-chairs: Marcus Fridén (Sweden) / Bonnie Murphy (US)**

200 delegates welcomed from 49 countries. Officially opened by Bonnie Murphy from US Bureau of Labor Statistics, co-chair of the Voorburg group. Trying new format for the meeting with sessions Tuesday/Thursday in next two weeks. Survey will be provided to delegates to provide feedback on the format. Meeting will be recorded to compile minutes.   
How to participate will be discussed on a slide before all of the presentations. Time will try to be maintained but good discussions may take place so it will be made flexible. Extra time has been provided for discussion.

Thanks to Statistics Canada and virtual planning committee.

2 vacancies to fill to prepare for 2023 meeting. Email Marcus/Bonnie with any nominations by Tuesday meeting next week. They will be announced at Thursday’s meeting next week.

**Welcome Remarks by Statistics Canada  
Chief Statistician of Canada, Anil Arora**

Video played from Anil Arora.

**Presentation of results from the Voorburg Group Alternative Data Survey**  
**Dragos Ifrim (Canada), Rohan Draper (Denmark)**

Dragos highlighted the increased role that alternative data plays in our processes over time and the reason this questionnaire was developed. The results of the Alternative Data Survey will not only allow us to have a global view of alternative data use but will also point out potential areas of development by industry, in our own countries. Only 16 countries out of 50 have responded to the questionnaire at this point and it is important to increase the response rate and implicitly, the representativity of results. Member countries were encouraged to contact Dragos and/or Rohan with any questions they might have and to send in the completed survey as soon as possible.

***Rohan Draper: Presentation***

* Thanks to all who have contributed to the survey data. 16 responses now and 4 more on the way. Please send in your material, even this week if it is possible. Any data received before the 22nd will be incorporated into the final dataset at the close of the Voorburg meeting.
* Purpose of the survey:
  + Offers the potential for advice and support from other members exploring admin data. We will also be able to see the trends going forward (snapshot of now) and look at the trends later on. Good source of information, good data source to focus on areas of interest.
  + To improve decision-making regarding SPPI development activities
* What’s new?
  + Good visualisations of the results – finding that coverage has increased to 16 countries (please complete the data and send it through to Rohan).
* Category recap – structured vs. unstructured data, stratified by internal or external sources.
* CPI usage – predominance of CPI usage in transportation and storage. Other industries also see quite a lot of usage. In some cases, this is a dominant input to some industries. This will be an ongoing area where we can learn from each other. Also, how to derive weights in a meaningful way.
* Multisource usage – Fair number of respondents reporting that statistical institutions have reported using multiple sources of data to produce a particular index. How to help with quality and what to consider when using multiple sources.
* Visualization of Transportation and storage example – overview of predominant data sources per 4 digit code
* Alternate data sources – when you first open up the survey, on the completion sheet you have to complete it at the 4-digit level. Need to fully open up the survey using the + icon or the 1 or 2 buttons. If you need help or support, email Rohan or Dragos.
* If you would like to do some analysis yourself, there are nice ways that the data can be presented. Look at the data by country and by 4 digit code. Can see what countries have done work in a specific place to see who is producing the statistics, what they are doing with it, what sources are being used.
* What’s next: more visualizations and information on weights will be added. Feedback would also be helpful. Get in touch with the team with any feedback or updates that you would like in order to have them incorporated into the survey. As well, any feedback of using the data from the survey in any kind of analysis would also be helpful.

**Session:** **Revisited Sector Paper – ISIC 77.30 Renting and leasing of other machinery, equipment and tangible goods** **Bonnie Murphy (USBLS)**

* 5 presentations in this industry from 5 countries.
* Industry classifications ISIC, NACE, NAICS. NAICS does not align directly with ISIC and NACE. Three that align well and three that align partially with them.
  + Product classifications are very close, with some as exact matches.
  + NAPCS diverges slightly from CPC and CPA.
* Service provided: this industry includes a wide variety of equipment. Services excluded is that it does not include the rental of machinery or equipment with an operator, financial leasing (treated like a loan as the risk of ownership is associated with the lessee).
* All countries reported that the RLMET represents a small portion of overall economic activity. Top percentage was 1% of the total economic industry.
  + Canada reported moderate level of concentration as did France.
  + However, Sweden and Estonia reported that it is dominated by small enterprises.
* Rental and leasing may be preferred to purchasing equipment. Timely especially during economic downturns and times of uncertainty. It also provides access to up-to date equipment, reduced up-front capital expenditures, can adjust production levels.
* Methods/Issues: combination of administrative and survey data. Variation across the countries.
  + Canada: aggregate level is reported which makes it difficult to allocate at a provincial level. The same with using admin data, hard to allocate. Also issues when primary activity of the larger entity may differ from the establishment. Also, difficult maintaining an accurate business register – StatCan allows for revisions, though they are relatively small.
* SPPI Price determining characteristics: type of equipment rented, duration of rental, inclusion of associated services (transportation, fuel maintenance)
* Pricing methods: France reported average price per day, contract pricing for large customers. Hungary – average prices, contract pricing. Estonia – price from company websites.
* Measurement issues for SPPI: difficult to maintain an accurate business structure, significant activity in multiple industries, sampling challenging.
* Quality adjustment: changes to the machinery or equipment, changes in the rental services.
* Two types of changes that need to be quality adjusted. Need to adjust the rental goods prices as well as the type of service provided. Difficult to assign a change of value to the rental industry.
* Prices of new vs. existing leases. If new/existing leases were priced differently, turnover would also need to be measured.
  + Spain 2019 Paris – collect info at the end of each year between the differences in new vs. existing leases in order to adjust weights. Changes in the number of new vs. existing leases may also be seen as price changes.
* Evaluation of measurement: current methods are fulfilling the purpose and no challenges fulfilling the output. Some challenges with multi-activity business structures and multiple interpretations of classification guidelines. Estonia collects information online, though there are drawbacks with this method of collection.
* Potential for web-scraping but there are some challenges associated with that. Lack of information available on negotiated discounts and existing leases.
* International progress: primary sources of data are questionnaires and structural business statistics. See more countries collecting data on turnover rather than pricing collection for this code however three countries said they did not collect turnover revenue statistics in the administrative data questionnaire. Sweden, Japan, Spain asked to respond more in-depth about their collection practices.

*Discussion*

Sweden, Japan, Spain asked to respond more in-depth about their collection practices.

**Sweden**: Barbro von Hofsten: they do measure the turnover measurement for these sectors

**Japan**: unsure about the practices. Confirmed after the meeting that they survey and publish output statistics on a monthly and annual basis.

**Spain:** Collects data for division 77 in turnover.

**Session: Cross cutting topic (1) – Alternative data in SPPI on air passenger transportation (51.10)** **Nicolas Studer (France)**

* Will discuss the work in progress on the air passenger transportation.
* Will consider making changes to their practices
* Until now they have been using a proxy to collect data on the SPPI for this sector
  + Using CPI for air passenger transportation
* Differences CPI vs PPI
  + CPI takes into account imports (sales by non-resident airlines)
  + CPI does not include sales to business (BtoB) and exports (BtoE)
* Prospect: Using data collected by the Civil aviation authority
  + 200,000 prices collected each month by a bot using the digital booking platform Travelport for members of GDS.
  + Low-cost airlines collected through webscraping.
  + Generalized webscraping tested but provided less robust
    - More difficult – you have to get authorization from the airline, if not the website will deny access to the bot. Authorization can also be removed at any time
    - Websites are changing from time to time. When you invest in a bot to work on a structure of a website and the website changes you have to develop another bot.
    - Easier to directly connect on data from digital booking platform
  + Issue: these are prices for non-realized transactions
* Prices all taxes included for roundtrips with departure from France available at the time of the search
  + Includes only one luggage and credit card fees which is consistent with ISIC as other services (additional luggage, special seat or meals, travel agency fees) are in activities of travel agencies (63.30) and other supporting air transport activities (63.23)
  + By product: origin/final destination, correspondence, air carrier, class
    - Sampling based on traffic, eventually 377 itineraries from 60 companies covering 46 % of traffic from France
    - Weighting of the products is based on annual traffic data and MIDT (booking) data: allows to “correct” for connections
  + By profile: advance booking (2, 11 and 20 days, 1 1/2, 3 and 6 months); duration of stay (< 24 hours, 3 days, 1 and 3 weeks) flexibility (exchange and reimbursement without fees, lowest price)
    - Weighting of the profiles is based on annual passenger survey (ENPA): distribution computed for 9 different markets
  + Weighting of the products is based on annual traffic data and MIDT (booking) data
* Currently used to compute a Passenger Air Transport Price Index (PATPI): prices offered by resident and foreign airlines to resident passengers (both enterprises and households).
* Difficulties
  + What is a foreign airline? These are airlines we are not interested in.
    - Production should be done in France in order to be included. But the flag of the airline providing the service is not the most important: one need to know if the service is originating in France existence of a hub, place of residence of the aircrafts....
  + Distinguish between clients between enterprises and household
    - No data on travel purposes (private or professional)
    - Use the class as a proxy – not a very good proxy but it is the only one that can be thought of.
    - Use the data in the passenger surveys to adjust/calibrate the estimate based on the proxy: proportion of people travelling in business class actually doing it for business purposes or not.
  + Distinguish among clients resident and non-resident
    - Proxy of destination. There is an assumption that internal domestic flights will be mostly residents whereas non-resident will only fly on international flights.
    - Assumption that the price of the same round trip will be the same with departure abroad (not wrong but different seasonality).
      * I.e., France à New York = New York à France
  + Taxes
    - Taking off taxes to compute base price indices

*Discussion*

*Tan Bing Xiang (DOS):* Hi Nicholas, thanks for the sharing. We are just curious what bot are you using to collect the air fares? We would also like to check if air fares can be consistently collected from GDS (What I am trying to say here is that is there a situation where a specific airline can decide to block us from collecting their air fares via GDS)? Thanks a lot!

*Nicolas Response:* We are not collecting the data ourselves, but through another organization. What they did to circumvent this issue is dealing with GDS directly.

*Tan Bing Xiang (DOS):* Last I heard from the GDS, it was mentioned that it could be possible that airline can decide not to provide their air fares via GDS to a specific user.

*Nicolas Response*: Nicolas Studer to ask colleagues to get the answer. As explained to him they tried to do web scraping, but the airlines forbid them to get the data from the website. That’s why they chose this solution with the bot in order to get around the issues.

*Jean-Michel Zürcher (Switzerland)*: Do the flights considered for B2C and B2B have different characteristics (I am thinking of destinations, time spent on site, etc.)?

*Response*: Something they will be able to look at in the data on the passenger survey (ENPA). They collect all of this data on prices for different markets on profiles and then will look at the profile of the air passengers annual survey to see B2C and B2B. Will be able to look at the composition of the different flights by profile and product. The characteristic missing on the passenger survey is about how many days you book the flight before.

*Murphy, Bonnie – BLS*: You may have said this already but do the prices you collect include discounted fares or fares that are free (use of frequent flyer miles for example)?

*Response*: No it doesn’t collect this. This is an issue with collecting the prices in this method.

*Murphy, Bonnie – BLS*: US uses a practice to collect for this. Collect data directly from airlines, select origin/destination, ask the number of people on the plane (when tickets are issued) – give total revenue per flight and the number of people on the plane. Receive the data directly from the airlines in a sample but the bots do provide greater information but does not collect for these discounted/free fares. US does not use alternative data in this industry but looking at alternative dataset to see if they can use it.

*Dragos*: considering the changes from CPI to SPPI, does the movement line up for the airlines? Sometimes it is not worth the hassle to change when comparing the results as they align very much the same.

*Nicolas*: Difficult to know beforehand. When you have experience you can look at this but unsure of the answer at this time.

**Session: Continuation of IMF collaboration (final results paper)**  
**Andrew Baer (IMF)**

* Goal of the project
* Compare large private data set of hotel prices to official PPIs for hotel lodging services
* Japan, Canada, India, UK
* Research questions
  + Do price indexes based on the much larger sample size of properties and transactions produce different trends than official PPIs?
    - Looking at much larger data set from the alternative source
* If so, is there a consistent pattern of trend differences across countries?
* Would real output growth for lodging services look different if the price indexes based on the large data set were in place in the past?
* Can the differences provide any insights for measuring the benefits of sample size expansion against the costs of introducing unit value bias?
* The basics – matched models and unit values
  + General best practice for PPIs – matched models
    - Select a sample of well-specified goods and services, monitor their prices in future periods when they are sold under similar transaction terms.
    - Update sample periodically, perform QA when necessary.
  + Alternative approach – unit values
    - Collect the price for a good or service as the total sales value generated from all transactions divided by the number of units (quantity) sold.
    - Subject to unit value bias – likely to be affected by changes in the mix of types of transactions that are included each period.
    - Advantage – low-cost method to increase number of sample observations
* If data collection resources were unlimited and response burden unconstrained, we’d always advise matched model approach since it is not subject to unit value bias
  + Instead of looking at one room, one night, we would want to look at all of their rooms, every night of the month. But this would be hundreds of transactions. This would be ideal if we didn’t have to worry about response burden or burden of having that much data.
* In real world, the decision may be based on assessment of increased precision of larger sample weighed against loss of accuracy due to unit value basis
  + This is difficult to measure (any ideas from the Voorburg delegates?)
  + This study produces only anecdotal information about this trade-off
    - Looks at what can we learn about the differences of having a larger sample size
* STR Historical data
  + For this study, the IMF purchased historical monthly accommodations data for Canada, UK, India, and Japan from STR.
  + STR maintains a Census of all properties with 10 or more rooms in these countries
  + 40-70% of these properties report supply of available rooms, room nights booked, and accommodation revenue each month.
    - Price = revenue / room nights booked.
  + Information for non-reporting properties is estimated based on reported data for similar properties
  + No property-specific information is provided, only aggregations across all Census properties at the national level
* Benefits of increasing sample for accommodation services
  + Different price dynamics based on market demands
  + A large hotel may make as many as 30,000 distinct room night transactions in a month
  + NSOs typically collect only 2-6 transaction prices from a property
  + Dynamic pricing compounds problem - algorithmic models that continually estimate profit-maximizing price for each room night based on evolving supply and demand signals
* Channel shifting difficulties
  + When demand is low more rooms made available through global distribution services (at a lower rate) while keeping the website price the same
  + Unit value would pick up on changes through channel shifting
  + Increased quantity of lower price rooms through 3rd party seller
  + Would not be able to capture this in match model data
* Cost of unit value bias for accommodation services
  + Erroneous price changes shown if the occupancy rates of different types of rooms vary from month to month
    - Greater occupancy in suite rooms from one month to the next it would not be something we would want to reflect in the pricing – not pricing changing but higher occupancy in the higher priced rooms
    - Unit value may reflect change, even though all prices remain the same. This is really a volume change, as Hotel A is selling more suite nights in Month 2
* Sample size comparisons – number of properties
  + Sample size for all countries using the new data was significantly larger
* Results of analysis
  + There is no clear consistent pattern between countries, and the relationship between the big data indexes and the official PPIs is not consistent over time in the same country
* Canada
  + SPI vs. STR
    - STR consistently increasing at lower rate than the official PPI 2001-2017
  + Because STR coming in lower there would have been more real lodging output if the indicator had been used in this period
  + 2017-2019
    - Radical differences in the data sources
* UK SPPI for lodging vs STR
  + Opposite of Canada
  + SPPI showing more growth than the STR large dataset
  + Pattern starts reversing after 2015, though in the opposite way to Canada
* Japan SPPI and CPI for hotels vs STR
  + Data quite comparable between the two data sources
  + Japan collecting unit value prices but with less properties
* India CPI for Lodging vs. STR
  + A PPI for hotels is not yet available
  + The CPI includes lodging prices for resident households
  + Different trends than STR but hard to compare since scope is very different\
* Some additional questions/conclusions
  + The big data unit values may do a better job of picking up short-term fluctuations
    - Short month to month changes the bigger data source is more precise
    - Long-term the same similar trend might be seen
  + No clear evidence that the unbiased matched model indexes are inadequate for measuring long price trends – so maybe no reason to risk unit value bias?
  + But there are interesting (and completely contrasting!) differences in the price trends for big data and official PPIs in Canada and the UK
  + Other potential areas of study – air transportation?
    - Potentially more stability in the output measures

*Discussion*

*Dennis Matthijs:* Statistics Netherlands get big companies to send all their relevant data rather than sending them a questionnaire. Legislation allows them to ask for it and they then receive a bigger dataset.

*Response*: still a risk of bias in using this method. Would be interesting to look at specifics in a single property based on unit values or use the track match models.

*Angela Gast*: Business clients have specific usage of rooms. May miss effect of measuring B2B as measuring B2all. Difficulties defining what rooms are being collected. All the data collected may disperse the impact coming from B2B.

*Response*: It is less common to book through corporate travel agency, so difficult to capture business to business expenses as now people use their personal information to book their work trips and get reimbursed later.

*Angela*: Perhaps look at when the traveling is taking place to hypothesis and figure out what is business and what is leisure (weekdays vs. weekends).

*Bonnie Murphy*: Looking at the average prices – were they weighted based on the number of transactions when shifting for the Booking.com transactions?

*Response*: Only received a single average daily rate for matched models of properties over time. Average nightly rate each period by property.

*Bonnie Murphy*: Could you weight the number of transactions business vs. leisure would that offer a better option?

*Response*: You wouldn’t want to weight on this as you would want to show difference in buyers by the pricing.

*Bonnie Murphy*: Within each category could weight with the aggregators?

*Response*: Would be interesting to see one room over time when looking at the same service over time.

*Aizcorbe, Ana:* Hi Andrew.  Really interesting paper!!! Your "channel shifting "issue is related to the "outlet substitution bias" issue that has been studied in the context of Walmart vs grocery and taxis vs UBER, for example.  Marshall Reinsdorf wrote a paper with Diewert and Nakamura recently that might help you explain the different patterns across countries.

**Poster session – Experiences of modernising production systems** **Dragos Ifrim and Xin Ha (Canada)**

**Poster session – Experiences of modernising production systems** **Dragos Ifrim and Xin Ha (Canada)**

* Producer Prices Division of Statistics Canada is reviewing its production processes and practices. This involves:
  + Moving from SAS to R and Python
  + Simplifying processes, maximizing use of Admin data and minimizing manual intervention
  + Training staff, etc.
* Retail Sales Price Index (RSPI) was presented as an example of the work done
  + It incorporates scanner grocery (transaction) data from the CPI (approx. 1.5 million price quotes monthly)
  + It includes motor-vehicle sales from Admin data sources
  + Transparency has increased - All production code Is stored on GitLab (a local Git version) and available to anyone in StatsCan
  + Analysts are expected to be proficient in R/Python, have full access to both the process and data. A process exists to track changes to production code (if any)
* Previous production system
  + Complex process
  + EQ data: Survey respondents and admin data sources
  + Combined in the database, process of data, then fed to corporate system to calculate the price indexes
  + A team was present to do each specialized step in the pricing index
    - Was very difficult for analysts to know the range of transformations done to the data from the moment it was reported to StatsCan until the index was created.
  + Each program required separate IT processes and assistance
    - Extensive resources required to make changes to processes (for ex. to introduce a new calculation method)
* New vs. Old Process
  + Simpler ecosystem
  + Changes done without IT support
  + Easier to adapt
  + New processes can be added
  + Much faster, production runs in 5min vs. 4 hrs with the old process.
  + Uses open source
* Analysts have better understanding and control of the data which helps in the analysis process
* One team in control of the entire process, rather than multiple teams in charge of separate processes
  + Better awareness of what needs to approve and better analysis of the data
* Retail Services Price index: Demo
  + Project page is stored on Git-Lab – repository manager for project codes
    - Easier to edit, review, leave notes on code
    - One-stop shop to manage a project
    - Track outstanding tasks
  + Commit: allows for you to track edits to your code – unique ID that allows you to see what has changed in the code
    - If you don’t like the changes, you can revert to a previous version
  + Continuous development feature – will have warnings to ensure correct outputs are being produced
  + Analysts have more in-depth handle of the processes as they are the ones in charge of modifying the code
* Run the output in R and you will also have all of the scripts used which leads to reproducibility and tracking of how the process was created for that cycle
* Live demo can be done by requesting information from Statistics Canada

**Mini-presentation – Data center colocation (addition to ISIC 63.11 presented in Paris 2019)** **Dennis Matthijs (Netherlands)**

* Data center colocation is a young industry but gaining traction in the Netherlands
* Question: Are there any countries who have researched data center colocation?
  + No responses
* Currently do not have classification of their own, part of NACE 6311 and ISIC 6311, NAICS 18210
  + Related digital industries 58.2, 62, 63.1
* Challenging part of the economy for SPPI as it contains a wide range of different complex activities
* Last year spoke to businesses to gain an understanding of the industry and will begin collecting data soon
* Data center colocation
  + In addition, products have been coded to allow for more detailed analysis at product level indices when needed
    - Interruptions are significantly less likely using this strategy
  + Instead of placing their computer servers in their own space, more and more businesses choose to entrust them to an external data center and to work with them remotely.
  + This holds several benefits. Collective costs are lower due to economies of scale. But most importantly, data centers are designed to ensure that the IT equipment never fails to function. Buildings are heavily guarded and all critical infrastructure, such as the power supply and the cooling system, is redundantly installed, such that the risk of a disruption is near-zero.
  + IT equipment stored in cabinets within the data center locations
* High market concentration, only about 25 businesses
  + Account for 15-30% of turnover in NACE 63
* Google/Microsoft have huge data centers in the Netherlands
* Cloud vs. Colocation
  + Cloud is when customers rent IT equipment to work with remotely.
  + Colocation – still work with own IT equipment but rent the space to store the equipment
* Output
  + Turnover mostly consists of the monthly bills that customers pay
  + Bills consist of:
    - A fee for the space rental
    - A fee for the direct energy usage of the customer’s IT equipment
    - A fee for the cable rental (the customer’s IT equipment needs to be connected to the outside world through one or more cables, “patching”)
      * Customer needs to rent the cables from the data center to reach outside world
  + 5-10% of colocation turnover from set-up services – installation provided to the customer
  + Data usage of the customer’s IT equipment is usually not billed by the data center, but by internet providers (NACE 61). The data center only acts as an intermediary to ensure the customer can make use of the telecommunication services offered by internet providers.
    - Outside the scope for data center colocation
* Challenges when measuring turnover
  + Eliminating single-tenant data centers
    - Only one client, or for their own use – should be out of scope
  + Transfer prices of hyperscale data centers (especially in the Netherlands)
  + Client located abroad
  + Measuring transfer prices rather than regular market
  + Lack of a separate industry class for data center colocation
  + Discrepancy between turnover on the business level and turnover on the activity level
* Data center colocation prices
  + Leave high skills (Google/Microsoft) out of scope
  + SPPI focuses only on data spaces that rent out to various clients
  + Main quality aspects of the colocation service (the monthly bill part):
    - Location: the more “connectivity”, the higher the price
    - Space: measured in terms of rack units and full racks
    - Power: measured in terms of kW
    - Energy usage: measured in terms of kWh
    - Green energy or grey energy
      * Surcharge for green energy
    - Patching: number of cables and type of cables
  + Even if all of the aspects above are the same, customer A may pay a somewhat different monthly price than customer B
    - Bargaining power
    - Age of the contract
      * Length of contract could create price discrepancies between new and old contracts
  + Customized vs. Non-Customized services
    - Many small services and then few larger big contracts
* Methods
  + We have considered a unit value approach: collect the overall average price per kW, the overall average price per cable and the overall average turnover per kWh.
    - However, this felt risky, because some data centers will struggle to divide all turnover over components in the right way
  + Instead, we have chosen a combination of two methods:
    - Direct use of repeated service for non-customized colocation services
    - Contract pricing for customized colocation services
  + Non-customized services
    - We ask the respondent to describe several typical colocation services that many customers purchase and to estimate the average price that is currently charged across all current customers.
    - Problem: data centers struggle to do average price if there are different prices across same services based on the contracts
  + Customized services
    - We ask the respondent to describe several contracts and to submit the price that is charged to the customer, given a fictitious amount of energy usage.
    - Can’t ask for average price as only one customer per service
    - Benefit: services provided for a long time – can be followed over time
    - Caveats:
      * Customer may use different amount of energy each month (different bill)
        + Use fixed fictious amount of power
      * Customer may scale up or scale down (expected every few years)
        + Need quality adjustment based on the number of kW
      * Important to add new contracts over time into the sample
        + Market is growing over time and initial contracts will become outdated at some point

*Questions*

*Kel*: Hi Dennis, what is considered green energy in the Netherlands? Do their prices fluctuate more as compared to grey energy?

*Response*: May be a definition needed to be included. The data centers know the difference but Dennis does not know right now.

*Mathieu Thomassin (StatCan)*: Are you also working on the Cloud services? Cleaning businesses and identifying those that should be included?

*Response*: want to include it in the SPPI in the future, but not now. Colocation is currently bigger in the Netherlands compared to the Cloud usage. Hard to classify these companies into an industry. Other locations have tried using web scraping but experience so far has not worked well in identifying the companies.

*Mathieu Thomassin (StatCan)*:Will share results from Canada’s web scraping journey.

*Bonnie Murphy*: Netherlands and Canada did not have anything in the 2019 Paris paper. Maybe we could add to the paper using the administrative alternative sources data source. Could you update the paper?

*Bonnie Murphy*: Are the number of companies (excluding Microsoft/Google) are they concentrated or is it a bunch of small units?

*Response*: Around 25 businesses in total excluding M/G then 23. A few big ones and a few small ones not heavily concentrated with big companies.

*Question*: What do you expect the response rate to be? How long will it take you to release the survey information back?

*Response*: Usually SPPI survey around 50% response rate but for new survey will try to get more than 50%. About one month period to fill in the survey, may take some businesses a little longer. Usually about a half year-year before publishing the data.

*Bertha Rodriguez -INEI PERU*: Data centers colocation services are offered only by private sector? what is the role of the government.

*Response*: No role of the government. Private sector only.

**Discussion on new Voorburg mandate including imports, potential topics for the in-person 2023 Voorburg Group meeting**

* New mission of Voorburg Group
  + Realized there is no entity that addresses import price indexes for services
  + With approval from UN statistics bureau the mission statement has been extended to include this
  + Looking for more input for topics on import topics
* Need to plan the topics for the 2023 meeting next year
* 2 weeks ago document shared with potential topics for next year
* Industry, Methodology, Data Collection, Cross Cutting/Collaborative, National Accounts
* Provide new topics between now and next week
* Do not want to repeat topics where papers have already been written unless significant changes in the industry
* Email the committee, put them in the chat, submit your ideas
* PowerPoint presentation with ideas
* Looking for country volunteers for each topic
  + Preliminary volunteer basis, gain approvals from agencies on creating paper and presentations
* Plan is for in-person meeting for next year with a virtual option

# Day 2 – September 15, 2022

**Cross cutting topic (2) – Innovative price or output products that incorporate combined domestic and import prices or output in published data**

**Japan’s FD-ID price indexes: Incorporating import prices - Moegi Inoue (Japan)**

* The Final Demand-Intermediate Demand price indexes (FD-ID price indexes) are price index series by stage of demand that are comprised of the final demand index (FD index) and four intermediate demand indexes (ID indexes)
* By integrating prices of goods and services, and by dividing the intermediate stages in optimal manner in accordance with the Input-Output table, inflationary pressures in the entire Japanese economy can be tracked, including both goods and services sectors, and the process of price changes being transmitted from upstream to downstream stages in the production flow.
* 3 unique characteristics
  + Express the production flow by the FD index and four ID indexes
  + Each index is an aggregate index of all goods and services as inputs to each stage, so it has the characteristics of input price index
  + Incorporate import price indexes as well as producer price indexes and input prices to each demand stage
* Regarding the trends of the price indexes for Stages 1 to 4 of intermediate demand and the FD all commodities index (2015=100) for the period from 2021 onwards, the price index for Stage 1, which is the most upstream stage in the production flow, showed the largest rise, reflecting a rise in international commodity prices, among other factors. On the other hand, the price indexes for downstream stages and the FD index showed moderate rises, with the pace of rise becoming slower in more downstream stages.
* This indicates that the price shock in upstream stages have been absorbed through the production flow.
* In Japan, industries depend on imports for most of the raw materials needed for their production, and imports also play an important role in each stage of demand. Therefore, in order to examine the inflationary trend and the transmission process of price changes, it is important to take into consideration the effects of not only domestic prices but also import prices on price changes in each stage of demand.
* The price changes in Stage 1 are attributable almost entirely to the price volatility of import goods. This may imply that price changes in Stage 1 are caused by such factors as international commodity prices, overseas economic conditions, and foreign exchange volatility, rather than by the domestic supply demand factor
* However, in more downstream stages (Stages 2 and 3) in the production flow, the contribution of prices of import goods becomes smaller while the contribution of domestic goods becomes larger.
* In Stage 4 and FD, not only is the contribution of domestic goods large, but also the contribution of services, which is unremarkable in upstream stages, becomes clearly visible.
* The results of these breakdown indicate that price changes for import goods, which represent an exogenous factor, are important for Japan in upstream stages of intermediate demand in the production flow but that the price changes are absorbed stage by stage as the production process moves on.
* Scope of the FD-ID Price Indexes
  + The FD-ID price indexes cover both of goods and services transactions, including transactions not only in domestics (domestically produced goods and services) but also in imports.
  + If we look at the scope of the FD-ID price indexes in reference to the I-O table, we see that those indexes are distinctive in that goods and services sectors are classified by stage of demand in the flow of production (column sectors of the I-O table), rather than by type of goods and services (row sectors of the I-O table).
* Stage Assignments: Overview
  + In order to divide intermediate demand into four stages, the value of inputs to goods and services sectors is used based on the I-O table (2015 base).
  + Each domestic sector is classified into four stages using cut off variables (X=70, Y=65, and Z=60).
    - Based on 2015 table
  + The cut off variables are determined through three processes; (1) provisional assignment of sectors to stages, (2) additional optimization, (3) judgment-based adjustment (For more details, please refer to Appendix 2-5).
* Stage assignments: Major Sectors
  + As a result of stage assignments, each sector (good and service) is classified in each stage.
  + Major sectors below show that these are appropriately classified in accordance with the flow of production, which can capture the supply chain structure.
  + Stage 1
    - Crude petroleum
    - Petroleum products
    - Crude steel
    - Worker dispatching Service
    - Wholesale trade (Building materials, Minerals & metals, etc.
  + Stage 2
    - Plastic products
    - Steel products
    - Financial service
    - Advertising services
    - Road transport
    - Internet-based services
  + Stage 3
    - Motor vehicle parts and accessories
    - Integrated circuits
    - Liquid crystal panel
    - Air transport
    - Wholesale (Machinery and equipment)
  + Stage 4.
    - Soft drinks
    - Passenger motor
    - Machine tools
    - Personal Computers
    - Hotels
* Calculation of weights for FD-ID Price Index
  + The value of inputs to goods and services sectors (the 2015 I-O table) is also used in the calculation of weights.
  + For the ID indexes, the value of internal flow is excluded from the weight calculation in order to avoid the “multiple counting problem”
  + This is not the case with imports because imports are not produced in the domestic production flow and thus never counted multiple times
  + Example: In the Stage 2, There are ”hot rolled steel”, “cold-finished steel”, “Shearing and slitting steel products”
    - If we calculate the indexes without any adjustments, weights of these indexes become large. As a result, the effects of prices of steel products tend to be overrepresented.
      * Avoid the “multiple counting problem” by excluding the weights of trade within the same stage
* Matching of Price Data (Commodity-level indexes)
  + Commodity-level indexes of PPI, IPI (Import Price Index), EPI (Export Price Index), and SPPI that correspond to the row sectors in the I-O table are used as price.
  + The CPI is also used in the case of services for personal consumption (B to C transaction) since Japan’s SPPI only covers B to B transaction
* Coverage of the FD-ID Price Indexes
  + If there is no appropriate price index, these weights are considered zero and excluded in the FD-ID price indexes.
  + However, the coverage of the FD-ID price indexes as a whole is approximately 70 percent, which is generally sufficient level for an aggregation price index based on producer prices that integrates price of goods and services.
* FD-ID price indexes by goods and services
  + In addition to the FD-ID price indexes by domestics and imports, those by goods and services are also beneficial for capturing inflationary pressures especially when energy prices are rising/falling greatly. (Like recently)
  + Looking at decomposition of contributions by type of goods/services to year-on-year changes, although energy made the greatest contributions in upstream stages, such as Stages 1 and 2, its effects declined in downstream stages, such as Stage 4.
  + On the other hand, services made progressively larger contributions during the production flow, with the contributions becoming larger in downstream stages, including Stage 4 and the FD (final demand) stage
* Comparison of the FD-ID price indexes with existing price indexes
  + The FD index (domestic goods) is less volatile than the CGPI all commodities index, particularly when the energy price is falling/rising. CGPI all commodities index significantly overstates price changes for energy because it has “multiple counting problem”
  + FD index is considered more appropriate as an aggregate price index that represents the macro-level supply-demand condition in the stage of final demand in Japan because it covers goods and services in the final demand.
  + The FD index (excluding exports) has the scope closest to a scope of the deflator for domestic demand and two indicators follow mostly similar trends.
* Index series of the FD-ID Price Indexes
  + In order to meet many kinds of analysis, 109 indexes classified (i) by type of goods/services and (ii) by domestics/imports with respect to each of the ID indexes for Stage 1 to Stage 4 and the FD index are published on a monthly basis.
  + Three other special indexes, the FD goods indexes published within our old structure, are also continued.
* Publication of the FD-ID price indexes
  + FD-ID price indexes are published monthly for the data since Jan-2015.
  + In principle, figures are released on the Bank of Japan’s website on the 20th business day of the month following the reference month.
* Conclusion
  + The FD-ID price indexes can track inflationary pressures in the entire Japanese economy, including both goods and services sectors, and the process of price changes being transmitted from upstream to downstream stages in the production flow.
  + Japan’s FD-ID price indexes incorporate import prices because industries depend on imports for most of the raw materials needed for their production. It is important to take into consideration the effects of not only domestic prices but also import prices on price changes in each stage of demand
  + The Bank of Japan publishes 112 price indexes by domestics and imports, and by goods and services so that analysts can choose their appropriate indexes that meet their analysis purposes

**Inputs to Industry Price Indexes -** **Jon Weinhagen (US BLS)**

* Industry output indexes
  + Measure the average change in prices for outputs produced by domestic industries
* Input to industry indexes
  + measure the average change in prices for inputs consumed by domestic industries
* Why are they useful?
  + Industry level analysis
  + Price transmission analysis
    - Input and output for given industry to see how it impacts price
  + Deflation
  + Contract price adjustment
* Background information
  + Indexes have been calculated since 1986: inputs to construction industries indexes
    - Didn’t include wages, imports
  + 2015: first change to the series
    - Coverage expanded to select mining and manufacturing industries and the addition of service inputs
    - A lot of outreach done with feedback
    - Need to include imports
  + 2020: Introduced satellite indexes
    - Improved coverage
* Source data
  + BLS PPI commodity indexes to measure price change for domestically produced industry inputs
  + BLS NAICS-based import indexes (MPIs) to measure price change for imported industry inputs
  + Combined together
* Publication
  + Introduced new set of input indexes experimentally with the release of August 2020 data
  + Available for most 3-digit NAICS categories ν Measure price change for both domestically produced and imported inputs purchased by the NAICS grouping
  + Available on the PPI webpage
  + Updated monthly after both PPI and IPP release data
* Methodology: Product selection
  + Bureau of Economic Analysis (BEA) Input-Output (I-O) “Use of Commodities by Industries” table used to determine inputs to an industry
  + Shows set of commodities that an industry is consuming
  + Example: IO Industry, commodity, use value that industry is consuming
  + Map the codes to MPI to PPI data in order to measure price changes of imported and exported materials
  + Example: Tire manufacturing mapping directly from IO code to MPI title to PPI title
  + Example: Legal services: No direct map to MPI
* Methodology: weights
  + After the set of domestically-produced and imported commodities consumed by the industry are determined, BLS develops weights for each PPI and MPI index included in the industry input index
  + The weight for a given commodity within the industry reflects the relative share of the commodity in relation to total industry inputs
  + Weights are derived from:
    - BEA’s “Use of Commodities by Industries”
    - Census’ wherever-made (WEM) value of shipments (VOS) data
      * Gets the value of domestic production values in a given year
  + Methodology: net weights
    - As a final step, weights are converted to net weights by multiplying the gross weights by net input ratios
    - Net input ratios reflect the portion of the value of commodity that is produced outside of the industry
    - Applying net output ratio eliminates multiple counting by removing intra-industry transactions
  + Weight example: Inputs to transportation equipment manufacturing
    - Domestically produced inputs 70% of index, imported inputs 30%
      * PPIs included: power, tires, legal services, etc.
      * MPIS included: tire manufacturing, etc.
* Publication structure
  + General into specific structure
* Use example: industry cost analysis
  + Example Inputs to 336, transportation equipment manufacturing, excluding capital investment and labor
  + January 2020 – July 2022
  + Trend line of total, then broken up by domestic and imported inputs
* Use example: industry cost analysis
  + Can also break down domestic inputs further into domestic goods and domestic services
  + January 2020 – July 2022
* Use example: price transmission
  + Example: NAICS 316: Leather and applied products manufacturing, index levels
  + Input price index and output price index
  + Trend similarly
* Use example: deflation
  + Using the index analysis for social assistance (no output index)
* Index exclusions
  + Excludes labor and capital investment
  + Excludes imported services
  + Excludes any domestically produced services that BLS does not publish PPIs for

*Discussion*

*Question*: Are you changing the weights?

*Response*: Yes every 5 years with new census data.

*Ha, Xin (StatCan)*: Thank you Weinhagen, Jonathan - BLS! Question about data source for domestic inputs. It says the source of data is the PPI commodity indices. Aren't the PPI commodity indices based on an output price definition, the prices received by producers and not the prices paid by purchasers? If so, are there any adjustments done to these PPI commodity indices before aggregating to arrive at an industry level or are you using it as a proxy for input prices?

*Response*: Yes output indexes used as a proxy for true index. Take the index for a product and then take into account other factors in order to create an adjusted proxy measure.

*Question*: In France there is a huge price increase in the value change with prices being increased more for output prices than input prices. Input prices are not published but may also be a sensitive issue. Are you facing the same sensitivity in the US? Don’t you feel some consequences for this, because it can be confusing when you see output prices for a branch increase more than input prices. Methodology is not the same and when you only release aggregate product code you may not know what specific product is used in the industry.

*Response*: Have not had any complaints on it yet, working on keeping it accurate. Sound methodology so no real concern on this.

**Industry presentations – ISIC 85 Education Session leader: Christian Puchter (Austria)**

* Three presenters: two for output and one for prices (SPPI)
* Industry Classifications
  + ISIC classifications (United Nations)
    - NAICS (Canada, United States, Mexico)
    - Mexico and Canada have slightly modified NAICS
    - NACE (European industry classification)
  + NAICS more in-depth classifications than ISIC and NACE
* Product classifications
  + CPC (ISIC)
  + NAPCS (NAICS)
  + CPA (NACE)
  + CPC and CPA are very much aligned, NAPCS slightly different

**Turnover and output for Education (ISIC 85) in Sweden: Barbro von Hofsten (Sweden)**

* Market for education in Sweden
  + Mainly conducted in the public sector
  + “The independent school reform” in 1992 allowed education in market producing enterprises
  + Increasing number of market producers since mid 1990s gradually
* Basic data from SBS
  + Development of the proportion of students in market producing education enterprises according to the Swedish National Agency for Education
    - Development of students
    - Showing development of primary schools (lower proportion of students, 15% in 2020)
    - In pre-primary education 20% of the children attended private preschools in 2020
    - Highest proportion of private schools in upper secondary education (28% in 2020)
* Basic data from SBS
  + Market for education in Sweden
    - Market producers in education contributed with
      * 2.6 % of enterprises and employees
      * 0.8 % of net turnover
      * 1.6 % of value added in the total business sector 2020
  + Basic data
    - Just over 30,000 enterprises included
  + Number of employees per size class (SBS 2020)
    - 0-9 employee 20%
    - 10-19 employees 11%
    - 20-49 employees 18%
    - 50-249 employees (medium-sized) 21%
    - 250+ (25 enterprises) 30% of total employees
  + Development of employees
    - Biggest increases in 2011 and 2016
    - 2017 only year with a decline
    - 41% increase from 2010 to 2020
* Development 2010-2020
  + Net turnover and value added increased by 56 and 67 percent respectively from 2010 to 2020
  + 2015 good year for both value added and net turnover
* Turnover by product
  + Largest turnover generated by primary education activities (6-15 years of age) generating 26.7% (largest since 2013)
  + Then Upper secondary (21.1%) and pre-primary (19.2%)
  + Small percentage coming from other activities (2.6%)
* Turnover by customer
  + 73% is from the public sector
  + Businesses 14%
  + Consumers 13%
* Financing of schools/political debate
  + Financing of primary and upper secondary schools via a special” school fee”
    - Paid according to number of pupils
  + Primary-, secondary- and higher education free of charge
    - Because of the system with school fees
  + Political debate
    - Discussions about the limitation of market producers to gain big profits
    - Private schools distributing profits to shareholders
  + Further regulation of the school market?

**Turnover / output statistics: Ramon Bravo (Mexico)**

* Descriptions and characteristics of the industry
  + NAICS: The Educational Services comprises establishments that provide instruction and training in a wide variety of subjects. This instruction and training are provided by specialized establishments, such as schools, colleges, universities, and training centers. These establishments may be privately owned and operated for profit or not for profit, or they may be publicly owned and operated.
  + According to the results from latest economic sector, growth seen in establishments, personnel and income between 2013-2018
* Monthly services survey
  + Shows the evolution of activity in the short term
  + Income returned to pre-pandemic levels
  + Personnel not back to pre-pandemic levels
  + Only private schools surveyed in this
* Specific characteristics of the industry
  + According to the National Association of Universities and Professional Schools of Mexico, the impact of higher education in economic growth has not managed to have the expected positive effects on Mexico´s economic growth, because the productivity structure has been concentrated in service Sectors, which do not generate economic push forces.
* General framework
  + The general framework is integrated by the Statistical Business Register of Mexico (RENEM by its acronym in Spanish)
* Measurement issues
  + Inco sources: economic census (every 5 years), national economic surveys (monthly and annually)
* Evaluation of the measurement
  + Sampling Scheme Due to the very particular characteristics of each domain, there are different criteria for defining the sampling scheme, forming 5 groups, 3 of them with a non-Probabilistic design and 2 with a Probabilistic design.
* Statistical Precision Indicators
  + Statistics instruments that allow to determine the levels of reliability of statistics generated from surveys with a probabilistic design.
  + Standard error 🡪 confidence interval 🡪 coefficient of variation
* Conclusions
  + With data from the Ministry of Public Education, 88.6% of the school population attends public schools, while 11.4% corresponds to the private sector.
  + Private institutions of basic to higher education reported that derived from the COVID-19 pandemic and online courses, just over 10 percent of enrollment, equivalent to 398 thousand students, left their studies unfinished or did not enroll.
  + The above is just one example of the importance of having statistical information that allows for a correct decision-making in terms of evaluation or implementation of public policies, in that sense, it is relevant to integrate into the design of surveys new methods for the compilation of data, such as the exploitation of administrative records and / or agreements with associations of educational institutions, likewise, continue updating the frameworks and samples, since the private education sector in Mexico is very dynamic.

**U.S. Export Postsecondary Education Price Indexes: Hayden Swegal (US BLS)**

* *Post-Secondary Price index no longer calculated in the USA – cut in 2008 due to budget constraints*
* Definition of the industry
  + ISIC 8530 Higher Education
  + ISI 6110 – Educational services – BEA International Transaction Accounts
    - ISI closely follows NAICS
  + BLS – International Services Priced Indexes
    - The price indexes aggregation structures were based on the National Center for Education Statistics’ Integrated Postsecondary Education Data categories, often referred to as IPEDS
  + Definition of the Industry – BLS
    - Published the XPIs (Export Price Indexes) under four different trees/aggregation structures based on Room & Board vs. Tuition & Fees, Undergraduate vs. Graduate, Public vs. Private, and the different U.S. regions
* Market Conditions
  + Growing service sector excluding the COVID 19 impact in 2020 and 2021
  + Growth in education related accounts, accounting for higher share every year
  + In 2021, China dominated the export education in the U.S. on both expenditures and student counts
  + Students counted in here hold visas for education in the US
  + In 2021, foreign students in the United States pursued 1,142,352 degrees in higher education: 69,495 (6%) Associate, 443,440 (39%) Bachelor’s, 438,390 (38%) Master’s, 191,027 (17%) Doctoral.
  + There were 49,630 foreign students for K-12 education in 2021 with roughly 10% in primary school, while the remainder were in secondary school (grades 9-12). China sent the most students with 17,375.
  + The main providers of educational services included over 6,000 postsecondary institutions, 98,000 public schools, and 30,000 private schools in 2020.
* Measurement of XPI
  + General Framework
    - The Export Postsecondary Education Price Index was intended for use in BEA’s export education account to deflate the Net Exports portion of GDP.
    - Used by the BEA as the primary user
  + Measurement Issues
    - The data source: U.S. Department of Education National Center for Education Statistics’ Integrated Postsecondary Education Data System (IPEDS).
      * No sampling as must report into this data set
    - Administrative data source: IPEDS covers the universe of colleges and universities in the United States.
    - Pricing data: prices for tuition fees, room, and/or board for a full academic year by a full-time student.
    - Weights: calculated by multiplying the number of foreign students by the prices of their respective institution.
  + Methodology
    - Using IPEDS pricing information – assumed foreign students would pay what institutions charged out-of-state students
    - ExEd Unit – weights for the four trees = the 4 different classification possibilities \* Total foreign students at a given institute
    - Example: For example, if Studyhard University charges $10,500 for out-of-state tuition & fees for its undergraduates and it educates 75 foreign undergraduates, then the ExEd unit weight for that university’s undergraduate tuition & fees would be: $10,500 \* 75 = $787,500
  + Estimation
    - Imputation of some units based on same cell calculations (based on similar institutions)
    - If a company never reports a price, uses the mean of the group as a proxy
    - Publication was to be done at the beginning of June, first published in 2007
  + Methodology, Out of Scope
    - Some schools were considered out-of-scope for XPI.
      * Military and theological schools – fees likely covered by the government or by an entity other than the student
      * Schools that primarily provide “first professional” degrees, such as law or medical degrees – different cost structure
      * Schools that primarily provide online education – not considered within scope for XPI
      * Schools were excluded if offered outside the geographic boundaries of the United States and its territories
    - Prices/expenditures for a student’s books, local transportation, off-campus living expenses were not included because of a lack of data
* Evaluation of measurement
  + Benchmarking – comparing XPI against others
    - CPI College Tuition – index values
    - Common fund’s Higher Education Price Index (HEPI) – index values
    - IIE’s Open Doors – student counts, some prices
    - DHS - Student and Exchange Visitor Information System (SEVIS) – student counts
  + Potential improvements:
    - Calculation of superlative indexes, such as Fisher or Tornqvist
    - Revision of Export Education historical price indexes
      * Possibility of revision if necessary
    - Inclusion of additional education prices, such as books and living expenses
    - Inclusion of English as a second language, professional training, and other non-college/university postsecondary education options
    - Account for financial aid
    - Inclusion of first professional institutions
  + Second chance: mirror what BEA measures?
    - Comprehensive and robust secondary sources
    - Stick to postsecondary education, or cover all export education
    - Educational versus Immigration data (both statutory administrative data sources)
    - Possible use of NAICS 611 classification
    - In-house technical obstacles, start with publishing experimental indexes
      * Systems have had countless changes since estimates first produced 15 years ago
      * Experimental estimates to prove proof of concept may be used

**Discussant Remarks on ISIC 85 - Education: Erika Barrera (Chile)**

Good morning,

My name is Erika Barrera, and I work on National Accounts at the Central Bank of Chile

I will share with you some remarks about ISIC 85, Education, specially from the point of view of the national accounts. The purpose of my presentation is complementing the previous ones and motivating discussion.

First, I will make some highlights about the industry

I’ll talk about Output measurement from the perspective of the national accounts

I’ll mention some Challenges for measurement

And finally, I’ll suggest some Questions for discussion

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**(Highlights)**

Let’s begin

Some Important characteristics of Education Industry are:

* A wide diversity of institutional units: including schools, universities, further education colleges, vocational institutes, private driving schools, etc.
* Different institutional sectors: various levels of government, non-profit institutions and non-financial corporations.
* Coexistence of market or a non-market output (non-market = without economically significant prices).

-> Output indicators presented by Mexico and Sweden cover market output

* Oriented mainly to final consumption, but exports may be relevant. -> Price indexes presented by US BLS cover exports (postsecondary education)

Regarding the service:

* Education is an 'individual' service, delivered to students by educational establishments.
* And the teaching is generally undertaken in groups.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

In this context, an important question is how we define the output

* Education output is the amount of teaching received by the students for each type of education.
* Here is key the distinction between output and outcome. Outputs are services that result directly from the production process; outcomes are situations that consumers value, such as a high level of education, and this may or may not be a result of a production process and can be influenced by several external conditions different from teaching. National accounts deal with measures of outputs, not outcomes.

Considering the definition, we can refer to quantification

* The quantity of teaching received by students can be measured by the **number of hours** they spent at being taught. This measure is referred to as the number of 'student-hours' (or 'pupil-hours’).
* Where the number of hours is not available, the **number of students** can be an alternative (in case the hours of tuition that an average student receives remain stable over time).

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**(Output measurement: Market output)**

So let’s see how output is measured in the national accounts, according to international recommendations. There are important distinctions if we talk about market or non-market output

For market output - at Current prices, output is measured as income from services.

Regarding Volume measurement the preferred method is **deflating** output by suitable **PPIs** for each type of education service. The price indices should take account of the **quality** of the service provided and ensure that basic prices are used.

A second best method is using appropriate CPIs, corrected to valuation in basic prices and reflecting the quality of the service provided.

If these methods are not available, output indicators based on **'pupil hours**’ or **numbers of pupils** are acceptable. Using appropriate weights is important.

**(Output measurement: Non-Market output)**

For non-market output - at **Current prices**, output is measured as the Sum of the total costs of production (including intermediate consumption; compensation of employees; consumption of fixed capital; and taxes).

Regarding **Volume measurement,** since prices are not available, the possible alternatives are: direct output indicators; direct input indicators; input price deflation.

The preferred method is direct output indicator using 'pupil hours’. The use of numbers of pupils is appropriate for tertiary education and distance-learning. This criteria should be satisfied:

* Complete or nearly complete coverage of the indicators.
* An also, Weights for the indicators based on unit cost with an appropriate stratification

The least preferred Methods are one based on inputs; or a direct output method with incomplete coverage or insufficient stratification.

**(Challenges for measurement)**

There are many challenges in the measurement of Education Industry. Some of them are the next:

1. Methods selection depend on the availability of information and also specificities of the industry in each country.
2. It´s important accounting for quality (either by direct or indirect methods):
   * For this purpose it is key the distinction of services, modes of delivering teaching, and education programs.
   * When accounting for teaching quality, any external effect should be controlled.
3. New challenges arised from the Covid-19 pandemic: suspension of classes and movement to a remote learning environment.
   * methods and their suitability to reflect a decrease in the service had to be evaluated.
   * Discussion on quality differences between classes online and in person became more relevant.
   * In fact, in national accounts, some countries decided to implement changes in their methodologies or to introduce adjustments.

Questions for discussion

1. Any experiences measuring non-market output of Education on Service Statistics (other than national accounts)?
2. Any experiences compiling Education SPPIs? Any other experience compiling Education XPI?
3. Are there quality differences between education online and in person that should be considered?
4. How well output indicators reflected changes in the volume of service during the pick of the COVID-19 pandemics?
5. Any other question or comments?

*Discussion*

* Which were the effects of the COVID-19 pandemic in education? Please mention the data sources for private and public education.
  + Salaries provided during this time but services not necessarily provided to the children (Internet issues, etc.). Service provided was very poor.
  + Chile Response: Measuring output measuring direct output indicators. Measure the output using the number of students and adjust the figure using attendance, and complimentary surveys about the provision of online classes by private and public schools. Result was a severe decrease in the service in 2020.
  + Sweden response: private sector and public sector information available as well as survey on the non-profit sector. Pandemic impact, total number of students in universities increased because harder to get normal jobs and switch to online education very quickly.
  + Mexico: seen pandemic impact in monthly and annual surveys. Saw a decline in employment. Taking into account data provided by both of these surveys.
* Question for Sweden: Public and Private schools. Is the state paying the fees going to the private schools?
  + Yes this is correct. Based on the school fee it goes to either the public or the private schools.

**Cross cutting topic (4) – Dealing with economic shocks**

**Session leader: Christian Stock (Austria)**

* Papers dealing with economic shocks, mainly COVID 19 pandemic
* Effects of the COVID-19 pandemic and possible solutions (with analyses of business sector effects and new/adjusted surveys):
* Data collection
  + Delays of reported data (preliminary and final dissemination)
  + Sample size of surveys
  + Voluntary or mandatory survey
  + Business closures (often don’t know if it is temporary or permanent)
  + Contact channels to responding units
    - Reorganization of contact channels difficult
  + Home Office (NSOs and responding units)
  + Administrative data sources
  + Alternative data sources
* Dealing with economic shocks
  + Response rates
    - Decreasing response rates
      * Is it in general or due to pandemic? Or both?
    - Different strategies can be used
      * Increase of sample size due to lower response rates
      * Decrease sample size to reduce respondent burden
  + Treatment of non-response or missing data
    - Temporarily missing data
    - Imputation strategies in economic shocks

**Dealing with economic shocks: How Statistics Canada pivoted during the pandemic** **- Marie-Christine Bernard (Canada)**

* Statistical adjustments based on the economic shock in the pandemic in relation to the annual services industries program
  + Survey data collection during the pandemic
    - Streamlined approach for the Service Industries Program with a smaller portion of the sample collected, with a focus on critical units. Delay of several months, collecting data for reference year 2019.
      * Critical units identified by subject matter analyst
    - Challenges: many business closures, several new surveys related to COVID-19 at Statistic Canada to provide relevant data
      * Pandemic tracing was also done through the StatCan collection operations, so conflicting need for the services
  + Businesses closing January 2015 – December 2021 (see graph)
    - Many businesses closed in May of 2020, hard to have collection when many businesses closed
    - Delay in collection into summer of 2020 though normally would have happened much earlier
  + Weighted response rates for selected annual surveys of service industries
    - Stable up to 2018, but drop in 2019
    - Response rates for 2020 picked up but remain lower than observed in 2018
    - Right now in survey cycle collecting data for 2021 – trend of lower response rates is persisting
    - Hybrid and working from home models have been making collection process more challenging
    - Response rates 5-10% below pre-pandemic levels
* Statistical operations and methods: strategies during the pandemic
  + Unprecedented shock to the economy
  + Make use of current information from similar reporting units
  + Revised imputation strategy based on the conditions
  + In 2019 and 2020 more imputation than normal
  + 2019 imputation strategy was not revised
  + 2020 – imputation strategy relying on historical data was no longer as reliable as public health conditions, changes to operations environment, wage subsidies made 2020 not comparable to previous years
  + Used imputation of similar reporting units instead of historical
* New data developments: advanced indicators
  + Advanced estimate of retail sales
    - Need more timely information with rise of e-commerce
    - Make use of early respondents
  + Advanced estimate of wholesale sales
    - Largely relies on early respondents as well
    - More disaggregated by subsectors
  + Flash GDP
    - Born out of the pandemic to answer the need for more timely data
    - Kept at aggregate level for now
    - Based on methodology that combines qualitative and quantitative information
    - Trade-off between accuracy and timeliness
    - Well used between key stakeholders
    - Discrepancies exist between flash and annual estimates, but not substantial
  + Looking into ways to disseminate at a more granular level
  + Pandemic accelerated the advanced indicator which had already been in the testing phase
  + Trade off of accuracy and timeliness
    - Open for revisions at later time
* Making greater use of administrative tax data
  + Relying on this data for early estimates – operating revenues, expenses, salaries
  + Financial data for service industry groupings are usually obtained through the annual business surveys
    - Results typically published 10-15 months after the reference year
  + Goods and services tax (GST) – value added tax by Canadian government bore by the final consumer
  + Payroll deduction files – operating expenses and payroll expenses
  + Advanced information available 6 months before final data released
  + Administrative data collected on calendar year basis but annual surveys collected on the fiscal year of the company
  + Imputation in both administrative data and survey data
  + Direction of growth very similar between the two methods
* Statistical measurement challenges
  + Data gaps – digital economy
    - E-commerce
    - Platform based activities
    - Importance of the Internet in business environment magnified by the pandemic conditions
    - Annual program surveys collect information on e-commerce but only about half are published due to quality issues and low response rate
  + Lack of a business frame
    - Misclassification of business activities and inability of frame to capture digital technology services
    - Cloud computing
    - Cryptocurrency mining
    - New characteristics were added to survey questionnaires to better measure new activities
  + Leveraging alternative data sources
    - Administrative tax data
    - External data sources
      * Such as dataset measuring platform-based activities for rental activities – characteristics difficult to obtain through normal surveys
    - Digital economy businesses recently started filing GST tax
* Final thoughts
  + Quick change during the pandemic while majority of employees worked from home during the health crisis
  + Data user needs continue to change during the pandemic
  + Measurement challenges
  + Need for timeliness/accuracy
    - Changing environment puts more pressure on quality and timeliness of information

**Dealing with economic shocks – Corona pandemic in Sweden - Mariah Nilsson (Sweden)**

* Pandemic in Sweden
  + Covid-19
  + No lockdown
    - Recommendations to work from home if able to do so
  + Social distancing
  + Limitations for public gatherings
    - Pandemic laws on theatres, restaurants, etc.
    - 8 people at most to attend an event
  + Governmental support to businesses
    - Short time work allowance, income subsidies
* Pandemic at Statistics Sweden
  + Sustain production
  + Work from home
    - New for pandemic conditions, never permitted before
  + Pro-active measures to reduce negative effects
    - Production statistics and methodology
    - A lot of data published on bankruptcies and life expectancies
  + Limited consequences
    - No large impact on any main products or statistics
* Data collection during the pandemic
  + Functioned as before
    - Businesses obligated to respond to the surveys
  + Cancelled the use of financial penalties for non-response
    - To give businesses a break during hard times
    - Led to smaller response rates
    - Large, important businesses had challenges responding as work from home challenges made contacting them difficult
  + Increased sample size
  + Not for the annual service surveys, instead increasing the response time
  + Admin data
    - Based on income declarations which cannot be refused by businesses
* Response rates and imputations
  + 85% response rate in admin data
    - Very slight decrease during pandemic but not major
    - Mean value imputation
  + Survey data response rate 82.1%
    - Obligation to respond, though some do choose not to
    - Pandemic response seen in both 2019 and 2020 years
    - Response rates higher for the service industries compared to production industries
      * Freight transport and restaurants
        + Small enterprises
        + Less persons employed in admin
      * Legal activities and engineering
        + Office based
        + Personnel in admin
        + Better understanding of the obligation
  + Decreasing response rates an increasing problem
  + Same imputation method as for the admin data (mean imputation)
  + Recommendation – use the same method during crisis if possible to avoid break in time series
* Pandemic effects in the business sector
  + Not as significant as feared
  + Smaller decrease than the financial crisis in 2008/2009
  + Net turnover -3 % in 2020 compared to financial crisis decline of 7%
  + Employees (FTE) -2.5 % in 2020, same decrease as in the financial crisis
  + Subsidies and grants
    - Short-time work allowance
    - Lowered employer contributions
    - Reduced income
  + All corona related subsidies had to be kept separate from other subsidies
  + Subsidies allowed businesses to continue paying their employees even though they didn’t have to work
* Industry analysis
  + Net turnover 2007-2020 for selected industries
    - Air transport, Accommodation, Travel services, libraries
    - Huge declines seen in 2020 across these industries in the net turnover
  + Net turnover, industries with positive development during the first year of the pandemic
    - Pharmaceuticals, other manufacturing industries, research and development, veterinarians
* Going forward
  + Will be interesting to see how second year of pandemic is reflected in the service business sectors and how growth rates will change

**India’s Experience through the Pandemic & Steps taken to mitigate measurement challenges - Rajesh Sharma (India)**

* Introduction
  + The pandemic was an exceptional economic shock causing significant disruptions in daily life and changes to way of life.
  + Strict lockdowns in different phases were imposed in India to control the transmission of COVID-19 :
    - Lockdown 1.0: 25 Mar, 2020 to 14 Apr, 2020
    - Lockdown 2.0: 15 April to 3 May 2020,
    - Lockdown 3.0: 4 May to 17 May 2020,
    - Lockdown 4.0: 18 May to 31 May 2020.
  + Combination of social exclusion, activity and mobility restrictions, steep income losses, and severely weakened consumer confidence.
  + Marked contraction in private consumption demand, services output, and the labor market.
  + Contact based sectors such as tourism; restaurants, hotels etc. were severely affected, while others, such as agriculture and related industries, etc. have displayed a remarkable degree of resilience.
    - Some of these industries completely closed, even in times of partial reopening
* Economic impact of Covid-19 induced lockdown
  + Every economic activity was brought to a standstill with a collapse in both supply chains, and contraction in aggregate demand.
  + Countries operated below potential with a cumulative loss to GDP in initial months in 2020, while health risks persist.
  + India experienced a steep decline in GDP during the year 2020 which is said to be the biggest ever recorded for the nation.
    - Increased GDP in 2021
  + Major effects were felt on productive output of industries, workers' and communities' livelihoods, as well as national and international supply networks.
  + Manufacturing sector experienced the deepest drop, bearing the brunt of the output decline during the first wave of the pandemic in 2020
  + Largest decline seen in May 2020, with additional Decline in May 2021 (lockdown again)
* Policy Response in India
  + Relief packages with focus on the five pillars: economy, infrastructure, system, demography, demand
  + Developmental and regulatory policy measures announced for strong and robust revival of economic activities in the economy
  + Vaccinations made freely available, enabling gradual economy recovery
* Impact of Covid-19 on Survey based Price Indices
  + Data collection and its availability were severely hampered and maintaining the quality standards of statistics became challenging.
    - Entire country was in lockdown mode
    - Employees working from home and difficult communicating within the group
    - Learning to use strategies such as “WhatsApp”
    - Coping with ill employees
  + The compilation of indicators such as CPI, WPI, IIP, etc. were seriously hindered as a result of low response rates due to:
    - No contact with the selected units
    - Limited communication on data
    - Temporary shutdown of units
    - Quantum of imputation and estimation expanded with reduced response rate
      * Response rate formerly 75-95%, during lockdown dropped to 45-70%, post-lockdown it has increased again 75-95%
* Measurement Methods & Issues
  + Accuracy and reliability of several indicators faced the challenge of data inadequacy.
  + Data collection became more difficult due to the inability to conduct in‐ person visits and business shutdowns.
  + Authorities in charge of gathering and disseminating real sector statistics began exploring alternative methods of data collection and estimation of non-response.
  + Advice of multilateral agencies including the IMF and World Bank were sought on the practices to be followed for preparation of essential economic indicators in such challenging situations.
  + Traditional face-to-face interviews were replaced by telephone interviews and online surveys and other innovative methods.
* Wholesale Price Index (WPI)
  + Response rate declined to 65 percent, against the usual response rate of 93-95 percent
  + Many industry groups could not get complete data
  + All Commodities WPI could not be computed for the month of April-2020 due to unavailability of substantial data.
  + Out of the 22 National Industrial Classification (NIC) two-digit groups for Manufactured products, indices were compiled only for 5 manufacturing groups.
  + Price movement of Sub-groups/Groups of WPI was worked out by taking the prices of items for which at least 25% of price quotations were reported.
  + Operational factories/manufacturing units across the country were briefed about the submission of price data through different means i.e. through telephonic enquiry, email, data submission portals, etc.
  + The provisional figures of WPI (in April 2020) with limited response rate were made final after a period of eight weeks/ two months by ensuring submission of almost all the required price data using alternative methods.
  + Numbers of Virtual training/discussions were organized on a frequent basis with the Field Offices of National Statistical Office (NSO)
  + Price collectors were contacted through telephone daily for resolving their issues and duly validating the submitted data.
  + Manufacturing units were requested to submit the Expected/tender prices for maintaining continuity and smooth flow of price data during lockdown through different means.
* Consumer Price Index (CPI)
  + Price collection of CPI through personal visits of price collectors was suspended in March 2020
  + Prices were collected from 674 urban markets and 524 villages, out of 1114 urban markets and 1181 villages for commodities which were being transacted during the initial month of lockdown.
  + With limited transaction of products in the market, imputed figures for CPI for two consecutive months during the first wave of the pandemic in 2020 were published.
  + Partial indices of CPI were compiled.
    - No longer based on the personal visits, difficult to complete the focused survey
  + Field enumerators collected data through telephone from the designated outlets and augmented this with visits to the neighborhood shops for items being transacted.
  + A separate exercise based on the imputation of the missing price was adopted to address the non-availability of data for several sub-groups during lockdown
  + Price movement of Sub-groups/Groups of CPI was worked out by considering prices which met the principles of adequacy
* Index of Industrial Production (IIP)
  + Difficult time with aggregate method
  + Imputation method used was not accurate enough to continue during the pandemic as other measures were showing different results
  + Imputation methods changed during the pandemic
  + Assessed different methodologies
  + Resumed initial method of imputation once economic conditions back to normal
* Initiatives for other surveys conducted in the Economy
  + New high frequency indicators of economic activity were combined with conventional indicators to generate alternative aggregate measures that helped assess the impact of the pandemic on the economy and also the pace of normalization.
  + Electronic Data Submission Portal (EDSP) was modified with enhanced security features for facilitating International Banking Statistics (IBS) data submission.
  + Reserve Bank of India temporarily substituted its Computer Aided Personal Interview (CAPI) based surveys of households with telephonic surveys.
  + Training of reporting entities and investigators were also on online platforms.
  + Proportions of interviews verified through audio and telephonic verification were increased.
  + Alternative data sources used as a supplementary source
  + IT systems – able to make more robust during pandemic time
* Lessons Learned
  + Digitalization is very important and through digitalization we can get the data without much assistance
  + Pandemic provides smaller reaction time
  + Lot of effort to use the administrative sources of data
  + Pandemic has given the time to evolve the integration of alternative data with streamlined methods and systems

*Discussion*

* Is it mandatory under law to submit the survey in India and Canada?
  + Canada: It is mandatory to answer the business surveys
  + India: Mandatory under Data Collection Act 1978
* During the pandemic, in PERU information and communication (ICT) activity was one of the activities that grew the most overall in urban areas. In the other hand, informal sector increasing gave raise to bigger level of poverty. I think, breaking down the effects by urban and rural area is very useful. I wonder if India faces this problem?
  + Pandemic mainly impacting the urban areas. In rural areas, mostly fine as little economic activity such as travel occurs in the urban areas.
  + Little impact in the rural areas as agricultural industries not largely affected.
* 3 challenges in France: 1) Imputation – higher proportion of businesses gone bankrupt or closed altogether. Makes it hard to make imputation for non-respondents (selection bias). 2) Chain linking – updating weights annually, relying on a 2019 description of activity. 3) Seasonal adjustment – usual models may display outliers and during the COVID crisis this outlier definition is not very clear. Were these challenges also seen in the other countries?
  + Canada:
    - Imputation – a lot of uncertainty here because of closures, government help, etc. This did impact our strategy. We do accept that there will be revisions to the data once more data is received. Mainly using donor data to do the data as the historical imputation is no longer relevant.
  + Sweden: Not as much impact was held by the pandemic.
    - Keeping weights for 2021 as used in 2020
    - Seasonality – imputing with the year change not with the monthly change
  + India: Outlier treatment – challenges
    - Had to change imputation as other sources did not match with the data
    - Evaluated different results from changing the imputation strategy
    - Better to change the strategy as it aligns better with other data sources
    - Difficult to measure all of the closures

**Discussion on 2023 collaborative guidance paper on SPPI and output compilation during economic shocks including best practices for producing output and price indices during these periods.**

* Looking to have a collaborative paper for next year presented at the meeting
* Christian Stock to organize the collaborative guidance paper based on collaboratives for papers in previous years

# Day 3 – September 20, 2022

Cross-cutting Topic (5) – E-commerce

**E-commerce Session leader: Ruth Vizner (Israel)**

* How the pandemic changed the way that services are digitally provided
* Looking at country’s measurements of how these changes were implemented
* Measurements of e-commerce before, after, and during the pandemic
* How the pandemic transformed the way services are provided, ordered or delivered
* E-commerce measurement in US and Canada done through various surveys
* Challenges – definition, record keeping, distinction, under-reporting, volatility
* US Measurement
  + Redesigned e-commerce question in their survey in 2017
  + Change word “E-commerce” to “revenues from electronic sources”
  + From one question to three-part question
  + Made it better, but had a price because time series broken in 2017
  + Standardization of e-commerce in the Annual integrated economic survey
  + 2020 – to measure the pandemic and e-commerce activity, three questions added to small business pulse survey about e-commerce activity (how the pandemic affected e-commerce)
  + Business Trends and Outlook Survey (BTOS) – 2022 – E-commerce will remain one of the key concepts indices in the future
* Canada
  + Focus on measuring digital activities in international trade in services
  + Cross boarder digital transactions
  + STEC – trade by exporter and importer characteristics – services:
    - Share of services exports digitally delivered and profiles of exporters
    - Import transaction of services by businesses and consumers from non-resident enterprises
  + A new economic indicator of digital intensity in delivery of services exports
* Results
  + In both US and Canada e-commerce increased during and after the pandemic
  + Slows as the economy re-opens to in-person activities
  + Many questions remain for the future
    - How to measure
    - How will the pandemic continue impacting the e-commerce?
    - Who will benefit from the data?

**Suzanne Conard, Tristan St. Onge and Edward Watkins (US Census Bureau)**

* E-Commerce Measurement Challenges in the U.S. Service Sectors
  + Changing “E-Commerce” to “Revenues from Electronic Sources”
  + E-Commerce Standardization
  + Measuring Pandemic E-Commerce Activity
* Redesign of E-Commerce Question – 2017 Service Annual Survey (SAS)
  + A review of e-commerce data prior to 2017 Service Annual Survey (SAS) revealed the need for additional research. Among the observations:
    - Differences in definition and measurement of e-commerce across sectors
    - Company record keeping did not always reflect the distinction between electronic vs. non-electronic sales
    - Volatile year-to-year reporting at the micro-level, e.g., firms reporting zero dollars after reporting billions of dollars the year before
      * Differences reported even from the largest firms in the sample
    - Suspected underreporting
      * In comparison of comparing the figures to other countries and company information
  + Based on observations and cognitive testing, the e-commerce question was redesigned for the 2017 SAS:
    - Question wording was modified
    - Term “e-commerce” was replaced with “revenues from electronic sources (RES)”
      * Thinking that “e-commerce” term was confusing some respondents
    - Question was expanded from one to three parts to help distinguish the various ways respondents might record their sources of electronic revenue
  + Improvements realized from the redesign:
    - Revenue for e-commerce activity increased 70.6% between Survey Year 16 and Survey Year 17
    - The portion of e-commerce activity to total revenue was 4.2% in Survey Year 16 and 6.8% in Survey Year 17
    - Promising findings to correct the under-reporting of e-commerce in the country
  + Costs of the redesign:
    - E-commerce timeseries was broken in Survey Year 17
      * No comparability
    - New terminology may have introduced confusion among respondents
* E-Commerce vs. Revenues from Electronic Sources (RES)
  + Changes by sector from 2016 to 2017 cycles
  + E-commerce to Revenue from Electronic Sources
* Ongoing challenges around the collection and publication of e-commerce data:
  + Despite redesign, still experiencing some of the same challenges for review and analysis of the e-commerce information
  + Disparities in industry concentration
    - In concentrated industries where the aggregated industry total is heavily influenced by the accuracy and compliance from only a handful of companies, poor reporting has a greater impact on the reliability of the data
    - When large company fails to report, difficult to produce an estimate
  + Few opportunities to validate reporting
  + No administrative data
  + Limited additional sources of micro-level data to reconciliate
  + Infeasibility in tracking RES by NAICS/industry
  + Differences in how company keeps their data compared to how it is asked in the survey
* Revenues from Electronic Sources (RES) 2019 – 2020
  + There were no methodology nor collection changes to the SAS RES questionnaire during Survey Year 19 or Survey Year 20:
    - Preserving the time series is important
      * Important for stakeholders
      * Important after breaking it just a few years before
    - Insufficient time to conduct proper research or cognitive testing
    - Tweaking of question content did not align with standardization efforts across multiple annual surveys
  + Industry comparison of 2019-2020
    - Data users can compare revenue and revenue from electronic sources year over year from 2019 to 2020
    - Total revenue dropped in the services total but revenue from electronic sources increased in 2020
* E-Commerce Standardization
  + Services Timeseries Preservation
    - No current plan to adapt RES for new e-commerce environment:
      * Timeseries was already broken between Survey Year 2016 and Survey Year 2017
      * Stakeholders value year-to-year comparisons
      * Planning an E-commerce standardization effort across additional programs as part of the Annual Integrated Economic Survey (AIES)
        + E-commerce will be part of this plan
* Current Collection and Publication Activities – E-Commerce
  + Collection and publication revolve around the 4 major annual surveys
  + Looking at the e-commerce questions in each of these four sectors differs
  + Question wording different based on the type of business they are engaged in
  + Sampling methodologies also different – establishments vs. employer firms
  + Standardization will be done as transitioning to AIES
* Measuring Pandemic E-Commerce Activity
  + While the AIES transition is ongoing, pandemic measures were measured through many small business pulse surveys
  + Small business Pulse Survey (SBPS)
    - Experimental data product created as a response to the pandemic
    - Weekly survey to measure the effect of changing business conditions on the US small businesses during the pandemic
    - Conducted in 8 phases, each around 20 questions
    - Data collected commenced one month after declaration of the global pandemic
    - Timely information released
    - 2 years April 2020 – April 2022
  + Small Business Pulse Survey – Online Platforms
    - Since March 13, 2020, has there been an increase in this businesses’ use of online platforms to offer goods or services
  + Small Business Pulse Survey – Change in Business Practices
    - Comparing now to what was normal before March 13, 2020, has this business done any of the following?
      * Adopted or expanded use of digital technologies
        + Choice of this response was around 26-30%
  + Small Business Pulse Survey – Future Plans
    - In the next six months, do you think this business will need to do any of the following?
      * Develop online sales or websites
        + 12-14% of those responding in phases 2-3
        + 9-10% in phases 5-8
  + Small Business Pulse Survey – Publications
    - Shows how small businesses were dealing with the pandemic
  + Business Trends and Outlook Survey
    - Plan to implement survey from concept of small survey with more time to collect and test
    - Launched in July 2022 the Business Trends and Outlook Survey launched as the successor to the Small Business Pulse Survey
    - Created from SBPS lessons learned, stakeholder feedback, and leadership strategy
    - Continue design of content aimed at capturing economic baseline norms, unforeseen events, and recovery
    - E-commerce will remain as one of the key concepts indexed
* Questions and Considerations for the Future
  + Will the increase in business participation of e-commerce translate into better reporting?
  + Will the current confusion over e-commerce definitions and in-scope vs. out-of-scope activities lessen with the increase in e-commerce savviness?
    - Confusion from business to business within the same industry
  + What data will best inform stakeholders and data users? What would they find most useful?
    - Continue to get input to collect the most relevant and useful information
  + Are there new or emerging economic frontiers which would benefit from the collection and publication of e-commerce data?
  + How can we merge e-commerce business data with similar demographic data to present the most comprehensive understanding of e-commerce?
    - To see any trends that may emerge

**Enterprise surveys and the measurement of digital trade in services in Canada - Diana Wyman and Jason Aston (Canada)**

* Trade by Exporter and Importer Characteristics – Services
  + Digital Intensity in RY2019 (RY2020 release November 2022)
  + Not yet measuring inputs as not the same quality of information is available
  + Incremental steps being taking towards imports
* Non-resident platforms (Audio and Video streaming, Restaurant delivery, Personal transport) in International Trade in Services Statistics in RY2019 (advertising integration into RY2020 release November 2022)
* On the shoulders of Enterprise surveys
  + Enterprise surveys’ e-commerce and digital technology modules key to development of Canada’s International transactions in services Digital Module
  + Spotlight on
    - Survey of Digital Technology and Innovation (SDTIU)
    - Monthly Retail Trade Survey (MRTS)
    - Role they play to investigate concepts, key flags, etc. to allow digital team to build on their work to create module for annual services surveys
  + Anchored administrative data linkage for digital trade and digital intensity
* Digital intensity – Results RY2019
  + Trade by Exporter and Importer Characteristics – Services (STEC): Digital Intensity
  + New economic indicator of digital intensity in delivery of services exports
    - CODR dissemination of digital intensity of services exporters by MNE status, Employment size class and industry in December 2021 (12-10-0142 and -0146) to create STEC – DIGITAL
      * Publishing as much as possible, but a lot of analysis still ongoing
    - Research paper published in Latest Developments in December 2020
      * Question added to survey for Commercial Services
    - Result of NEW digital module into annual International Transactions in Services Survey (Commercial Services) now enhanced coverage to include exports of 25 services categories
      * Where data sources are available to dig into digital trade
      * Question first added in 2018 to look at digitally provided services
    - By adding services, received key details to see impact of delivering digital products
* Digital intensity: 36%
  + With $30 of $84 billion worth of exports were self-reported as digitally delivered in RY2019
  + Overall, nearly two-fifths of commercial services exports were reported to be digitally delivered in 2019, amounting to approximately 30.2 billion of business services. Typically, there are services provided through the use of platforms, applications and other means of transmission of materials online. This estimate of digital intensity provides the foundation for assessing enterprise-level adaptation over the course of the pandemic in 2020 and 2021.
  + SMEs account for larger share of commercial services exports but report lower digital intensity
* Foreign multinationals report higher digital intensity than Canadian multinationals
  + Non-US foreign MNEs operating in Canada reported the highest share of exports digitally delivered
* Digital intensity profiles across MNE and employment size profiles high digital intensity shares found in large MNEs
  + Looking at all services exporters, large enterprise, SMEs
* Non-resident platforms and enterprises representing imports into Canada
  + Audio and Video streaming, Gaming, Restaurant delivery, Personal transport) in International Trade in Services Statistics/Balance of Payments in RY2019
  + Integration of trade value between Canadians and non-resident platforms through consolidated value (leveraging horizontal collaboration)
  + Enhanced coverage leveraging administrative data (Goods and Services Tax, GST)
    - Reporting revenue generated in Canada even if physical location was not in Canada
  + Better understanding of players and how to capture/distinguish activity What’s Next?
  + Integration of trade value between Canadians and non-resident platforms – including data-driven advertising revenue RY2020
* Closing the data gap on non-resident enterprises generating revenue in Canada
  + For accommodation, ride-sharing and food delivery, it is necessary to break down the fee or margin from the overall revenue generated
    - Local restaurant delivery - domestic
  + For streaming (music, podcast, audiobook, video, video game) directly from non-resident enterprises via the non-resident streaming platform, imports tend to have a subscription fee that is channelled to the non-resident platform. This streaming fee flows from the household or firm in Canada directly to the non-resident firm, and therefore, the entire fee is included in the import value.
    - New data sources for subscription fees
  + For data- and advertising-driven platforms, data arriving from recent legislation, more complete reporting in annual financial reports by geographic segment and further mining of existing data sources can provide better estimated values.
  + For other digital infrastructure platforms, underlying payment and digital delivery facilitation mechanisms, domain name and security infrastructure, among other aspects of digital infrastructure, are necessary dimensions to capture to measure this activity.
* Enterprise surveys @ Statistics Canada and e-commerce
  + E-commerce and digital technology survey modules: The road to measuring digital trade in services
    - Survey of Digital Technology and Internet Use
      * Trying to capture enterprises engaged in e-commerce in early 2000’s
      * Vital to take these steps so in later 2010’s, we had some understanding, baselines, relations with the respondents to examine what was happening
    - Monthly Retail Trade Survey (MRTS)
      * Ensure timely measure of e-commerce
      * Monthly timeliness
      * Covid-19 pandemic response (relevant)
      * Digital technology content
      * Delineation between purchases and sales through digital ordering
      * RY2022 Delineation between digital ordering of goods and services digital delivery of services
      * Sample-based survey with high response (usually 90%)
      * When e-commerce added in 2012, good response
      * When pandemic started, able to confirm by mid-May what was happening in the sector as timely data
* COVID-19 impacts on retail e-commerce sales
  + Retail e-commerce sales nearly doubled during the first wave of the Covid-19 pandemic compared to January 2020
  + Closure of physical locations and lockdowns, limited hours, products not eligible for sale, unavailable products, inessential store closures
  + Gain an insight as to which sectors were benefiting to online world, which were quickly able to adapt to the digital environment
  + E-commerce sales slowed as the economy began to reopen to in-person activities
* Custom classification analysis
  + Shift to online sales in the total overall retail market in 2020
  + Permanent level shift to digital online sales – amount of activity gone online that will probably never go back to in-store
  + Online share of clothing and clothing accessories stores
    - Shopping malls heavily impacted by Canadian lockdown measures that were non-essential stores
    - Deeper dive by industry, could see larger share of online sales during the pandemic
* E-commerce customer locations
  + The vast majority of businesses with online sales (98%) had customers located in Canada, while around one-quarter exported outside of Canada
  + Just over one-fifth (22%) with online sales had customers in the United States
  + Just one-tenth (11%) of businesses with online sales reported having customers located in regions other than Canada or the United States
* Measuring Digital Intensity
  + Looking at number of webpages, percentage of Internet based sales, websites they are using to do the digital sales
* What’s next for Digital Intensity
  + Integration of value of non-resident digital platforms and enterprises into imports in RY2020 (release November 2022)
  + Feasibility of extending digital intensity estimates of enterprises to imports
  + Feasibility of extending digital intensity estimates of enterprises to exports/imports of goods and services

*Discussion on E-commerce*

* How do you calculate non-response rate and what was the disappearing enterprises treatment?
  + Can look into this response but does not manipulate the data.

**Session: Cross cutting topic (6) – Quality Change**

**A Proxy Approach to Quality Adjustment of A Service Industry - Rob Bucknall (UK)**

* Deflators
  + Allow for volume measure to be compared in real terms
    - Price changes removed from current price data series
  + Accounting for quality change is major challenge
    - Standard techniques not always suitable or practical
    - Consequently deflators of goods and services impacted by rapid quality advancement often tend towards upwards bias
  + Key aim of deflator challenge is to capture change in more timely manner
* Productivity within service sector
  + Low productivity growth in recent years for several service industries experiencing rapid tech change
  + Expect rapid tech change to lead to increased productivity
  + Productivity puzzle
    - Drop in productivity after 2008 crisis
    - Flatlined after the crisis not expected
  + Growing service sector in the UK
* Challenges in measuring service quality change
  + Adjusting for quality change of service is challenging due to the often heterogeneous nature of products
    - Tailored to client’s needs
    - Unique
    - Change from period to period
  + Quality of service is a function of its intangible characteristics
    - Reliability
    - Effectiveness
    - Customer satisfaction
* Pricing methods
  + Ideally the pricing method would inherently account for quality changes or productivity
  + Many of ONS’s SPPIs use time-based methods
    - Survey asks how long workers of different grades/positions work over given quarter
  + Model pricing would be more suitable but involves significant burden on respondent
* Case Study: Architecture & Engineering
  + Increasing use of emerging technologies over the last 10-20 years
  + Expect to see the impact of quality improvements reflected in the deflators
  + Existing deflators fail to recognise technological advancements and productivity revolution within the industry
  + CPA 71 dominated by small businesses (4 employees or less as of March 2021)
  + Quality change
    - Several technological changes to these services including 3D printing
    - Allows to positive implications for efficiencies
    - Current and expected use of technologies
      * Widespread use of new technologies
      * Highlights potential changes likely to take place over the next five years
  + Building Information Modelling (BIM)
    - In 2011 UK government commenced program encouraging use of BIM
    - Rapid use and awareness of BIM
    - Scale occurring since 2011 with increased use of BIM
    - Productivity influences from the technological advances such as BIM
  + Quality adjustment options
    - Implement a price adjustment using relevant proxies that could indicate a measure of quality change in the service
    - Use a pricing method which allows for inclusion of changes in quality or productivity, such as model pricing
  + Adjustment using satisfaction as proxy
    - Key Performance Indicators (KPIs) for the construction sector available from UK industry performance reports published by Glenigan
    - Satisfaction alone is a good indicator of quality as it likely related to many other indicators of quality
    - Upward trend of client satisfaction
    - Percentage change of satisfaction over time to calculate the index
    - Calculate the quality adjusted index as a ratio of the unadjusted SPPI to the quality index multiplied by 100
  + Results
    - Unadjusted PPI exhibits overall growth between 2010 and 2020, suggesting price increase
    - Quality adjusted SPPI exhibits flatter trends
    - By stripping out the quality improvements the price increases are offset
    - Expect resultant volume measures to show more growth than those calculated with unadjusted SPPI
    - Expect higher productivity growth over this period
* Conclusions
  + Adjusting for quality change of services is challenging
    - Heterogenous nature of services often prevents application of standard measures
  + Difficulties measuring quality change with increased digitalisation
  + Application of proxy approach led to plausible inflation rates which would lead to higher productivity estimates
  + We would like to continue our research by
    - Investigating additional sources which may be more suitable for developing quality indexes
    - Considering the use of additional KPIs such as people and environmental indicators
    - Investigating the potential of using this methodology for quality adjusting other service industries

Discussion

* Is it only for time-based industries or could it be used for other industries?
  + Mostly time-based industries as can be big burden on the respondent
  + Hoping to get feedback from Voorburg members to see if any other countries have approached the quality adjustments in any other ways
* How is satisfaction defined?
  + Sourced data from survey on the construction sector which covered engineering as well. Survey sent to companies in the industry and they produced the measures on satisfaction on how satisfied businesses are as a result of over time using the technologies
  + Rob to put link in the chat for the measures
* Do you price the construction surveys? Would you use the same strategy for this as for engineering and architecture?
  + Technologies used in other industries so potential to use the quality adjustment for other industries. Further research required but could go down that route.
* Is satisfaction survey a required survey or voluntary?
  + Voluntary. It is an industry report by a marketing company sourced for research.
* How do you deal with the subjectivity of satisfaction?
  + Proxy approach given a lot of research using the source data, but based on the quality of the data provided.
  + Scale of satisfaction transformed into percentages
  + Research has some weaknesses, but looking for feedback from other countries
* What is BIM?
  + Building Information Modeling – new technology which allows for increased efficiency within the construction sector for companies to improve their communication to more easily identify problems on the construction side before they start their project rather than beginning a project and finding problems later on.

**Presentation of participation in Ottawa meeting**

**Alignment of Methodology and Scope between Services Producer Price Indices (SPPIs) and Consumer Price Indices (CPIs): Developing a framework for using CPIs in SPPI calculation - Rohan Draper and Nicklas Elversøe (Denmark)**

* Foreword
  + This group task force was formed with the mandate of improving knowledge regarding SPPIs and CPIs
  + Usage of the alternative data survey – CPIs are the main alternative data sets being used to populate the SPPIs
    - Motivation to conduct this work and get the frame together
* A framework for using CPIs in SPPIs Recap – Trade Flows
  + Checklist of fitness of CPI data
  + Differences in valuation basis presents a significant risk to representativeness when using CPIs in SPPIs
* Main considerations
  + A framework for assessing the feasibility of CPI utilisation in SPPIs should include (not limited to):
    - Main purpose
    - Valuation principles
    - International trade
    - Nomenclature (concordance)
    - Product vs Industry
    - Supply and Use tables
  + Beyond these aspects there are other operation differences in the practical application
* Single-Source Data Collection “Traditional”
  + Direct data collection through questionnaires
  + Requires resources and time
* “Blending” – Using CPIs in combination with traditional data source
  + B2B (PPI)
  + B2C (CPI)
  + B2X (PPI/XPI)
  + Checklist for fitness of CPI data
    - Valuation principles
      * Need to clean it
    - International trade shares
    - Nomenclature (concordance)
    - Supply and Use (representativeness)
  + Low international trade shares and high household consumption is what you are looking for
* Sometimes an industry only needs to use the alternative data
* Questionnaire from 2021 – 15 countries supplied answers
  + Reported whether they use CPI and for what industries
  + Identified areas where CPI might be a good base for SPPIs
    - Prevalence of household consumption (B2C)
    - Typical industries identified
* Multi-Source data collection – A practical approach to “blending”
  + Once appropriate PPI candidates have been identified for coverage by CPIs and CPIs have themselves been deemed fit-for-purpose, explicit weights may be constructed for their inclusion in the aggregation structure of the PPI
  + The paper proposes a potential practical way of blending the indexes together. Predominantly this is based on the proportion of B2C and B2B shares of output with the supply use tables providing the price statistician with a sensible starting point
  + SUTs generally hold the information needed to create weights for blending the data sources together, but there are a few steps needed to calculate the output shares
* Are we doing this in an internationally comparable manner?
  + Using CPI in combination with other data sources and collected data
  + Need to have conversations with methodologists
* When using CPIs in the PPI some important factors are worth considering:
  + Communication – the need for closer communication between the CPI and PPI teams to ensure large changes in methodology and/or price movements are discussed and considered
  + Harmonization (general) – general principles and their application should be well understood and harmonized
  + Harmonization (per period) – ensuring good quality adjustment and imputation methods
* Feedback from Ottawa Group meeting, June 2022
  + Cooperation – need to cooperate, emphasized the growing importance of services. They were interested in hearing more about the practical application and use of SPI in NSIs today. Particularly where CPIs are being utilized as a proxy for B2B activities
  + Harmonization – interest in how CPI and PPI prices track each other over time and how valuation aspects can be controlled.
    - Stability (price)- markups can be very different over time
    - Quality adjustments – likely very different in a consumer versus a producer context
    - Product perception – firms vs. consumers
  + Standardizing data collection and cooperation between NSIs
  + General summary was that cooperation would be a worthwhile pursuit
* Next Steps
  + Positive step forward
  + More practical steps can be considered
  + Opportunity for test application and comments
  + Addition to SPPI Development Guidance
  + Follow-up with Ottawa Group
    - No negative thoughts towards this project
  + Official framework
  + Suggestions are welcome
  + Is the Bureau and Voorburg group satisfied that this taskforce is concluded successfully?
    - Is there interest in learning more about practical application?
    - Is there interest in further engaging with CPI forums?

**Presentation of results from Task Force on Outreach Karin Blix (Denmark), Riikka Korhonen (UK)**

* Outreach from the Voorburg Group
  + How to increase engagement within the Voorburg group
  + How to pass news within the group and from other groups with shared interest
  + Why?
    - To reduce the burden on the bureau members and their networks
    - To facilitate interaction and exchange of experiences between countries
  + How?
    - Introducing a newsletter
      * To be sent to other organizations and group members
      * To be placed at the Voorburg group website
      * Introducing a possible frame for quarterly publication
* Ideas for content
  + Get to know country/project
    - Get information from other NSIs or Voorburg members to hear about what they are doing
  + Notice board
    - Bureau news, member state news, international news, any other business
  + Events calendar
    - Upcoming meetings and webinars
* Call for collaboration
  + Voorburg members can showcase projects that they would like to discuss with other member states
    - Example: UK Deflator Development team
* Next steps
  + Information on projects that member states are working on and would like to discuss with other member states
  + How can the newsletter work in practice?
    - Rotation responsibility among members
    - Two or three with alternating responsibility and rolling change of members
    - Predefined task-list for the responsibilities to minimise the burden
  + When to start? Q4 2022 or better with Q1 2023?

**Global Network of Data Officers and Statisticians, and explanation of how you can create and maintain successfully a group on the Global Network - Alexander Loschky (UN), Yuxi Zhang (UN)**

* The Global Network of Data Officers and Statisticians
  + A facilitated professional online social community
  + Launched in October 2020 on Yammer
  + About 2400 users posted about 3100 messages and read about 250,000 times
  + Supports the work of all UN system entities
  + Improves communication and collaboration with NSOs
* Current demographics
  + 2400+ colleagues
  + Half from colleagues from NSSs and other governments
* What you can do on the network
  + Get informed about events and activities
  + Build capacities by attending live events
  + Advertise job and consultancies opportunities, call for papers or projects
  + Find and propose expertise on specific topics
  + Connect with global community
  + Exchange best practices
* How to join
  + Using Yammer Mobile Apps
  + Join by clicking link www.yammer.com/unstats
  + Ensure to use work email address not personal email address if possible
* When and why creating groups?
  + The global network
    - Providing a space where a restricted community can gather around a common challenge and start exchanging knowledge and directly useable information.
* When is it a good idea to create a group
  + When you want to address a specific topic
  + Exchange information or ask/receive feedback
  + Create/engage in a discussion
  + Engage peers to discuss it
  + Want to keep collaborating, having discussions between in-person meetings
* Conditions for a group to be active and useful to its members
  + Relevant audience
  + Clear objective – what topic will be covered?
  + Moderators allocate time to respond/curate content
  + Co-facilitators and champions actively involved
  + Regular and relevant information
  + Engaging discussions
  + Provide useful resources
* What types of groups exist?
  + Public groups – open to all
  + Private groups – restricted to members
    - Consider for the Voorburg group on the global network
* We can support you by offering:
  + Technical support on how to create the group
  + Provide information about what needs to be considered to make the group successful
  + Onboarding the Voorburg Group members
  + Provide advice and feedback on your launch strategy

**Vision for collaboration between meetings**

* What do we want to do in the future?
* Yammer not video but postings and discussion groups
* Newsletter and Yammer groups can be used as a prompt to get collaboration and discussion between meetings
* Looking for collaboration to allow for quicker information
* MS Teams vs. Yammer
  + Teams is good for collaboration within small teams working closely together
    - Virtual meetings
  + Yammer is more suited for larger groups – asynchronous messages to get feedback
    - Does not allow for virtual meetings, but can use it as a tool to plan a virtual meeting
* Does Yammer notify people that there has been activity within a group?
  + Yes – if things are posted and you are a member of the group, you will get a notification. Also a way to turn off the notifications.
  + Option for administrators to post “announcements” which are received by group members by email regardless of whether they had turned off their notifications.
* Poll in the meeting chat – having seen the Yammer presentation are you likely to use it to collaborate between meetings – 50 responses
  + 28% yes
  + 6% no
  + 66% maybe
* Would a quarterly newsletter be a good and viable idea?

**Classification topic – Wholesale trade Ruth Vizner (Israel)**

* Advance work
  + The Voorburg Group presents papers in 2008-2009 from US, Canada, Israel, Sweden, papers presented in 2010
  + Since then countries have updated their approaches recently Japan presented a preliminary research
  + Wholesale trade makes up a large proportion of output in many countries
  + It is important to develop and calculate turnover and SPPI for that industry
  + Because of the complexity in measuring, only few countries have SPPIs for that industry
* Classification issue
  + Wholesale trade is classified by products that are wholesaled rather than the service provided
  + This is an attempt to produce a collaborative classification suggesting services that might be appropriate for classification
    - Instead of product classification, using service classification
* Definition of the service
  + Wholesale output defined as the resale (without transformation) of goods from the manufacturer to another party
  + The wholesaling process is generally an intermediate step in the distribution of merchandise
  + Take the risk of the sale because they buy the products/services
* Classification
  + The different industry classifications are broken down into classes according to the type of goods traded
  + Split into groups according to the type of goods
  + According to the function and kind of product (food, household, machinery)
  + Detailed information and description of products and services
  + Distinction between wholesale trade on own account and on a fee/contract basis
  + Specialized wholesale goods
* Pricing method
  + The pricing methodology is the average margin price per unit – the difference between the selling price to the consumer and the acquisition price from the supplier
  + Commissions are the most common type of price in the Wholesale Trade Agents and Brokers industry
* Margin price
  + Difference between Selling price and Buying price
  + Can be the average prices or the unit prices
* Difficulties in Measuring
  + Measuring prices with constant quality
    - Should SPPI reflect changes in the quality of the good sold?
  + Lags between the purchase and sales of good
    - Large fluctuation of individual price
    - Negative prices
  + Import and export of goods
    - Margins affected by changes in exchange rates
  + Reporting complexity
    - Firms may consider data on margins to be sensitive information also complicated to report and cause respondent burden
  + Hard to measure this industry in a correct way
* Alternative classification
  + Need for more appropriate classification to serve for the best measurement of the industry
  + Suggesting services that might be appropriate for classification (Promotions, marketing packaging) instead of classification according to the type of goods traded

*Discussion*

* If all refrigerated goods require the same service then should they be grouped together? Bulk liquids very different service than wholesaling refrigerated goods.
* Unique way of thinking about it, unique take on a complicated issue.
* Were the respondents confused by product groupings? Would be interested to see this.
* How do you face imported inflation? Should you measure the price difference between import/exported goods?

**Poster session – Experiences of modernising production systems Andrew Carey (UK)**

* Look at journey of changes to the production systems
* An efficient modular system designed to evolve
  + The UK undertook a methods change in 2019/2020 transitioning from a rebased to a chain-linked series
  + Our legacy processing system was unable to facilitate this so we migrated to new platforms to modernise processes, improve performance and analytical tools as well as adopt online data collection tools
* Legacy system
  + Motivations to take up the new system as there were some shortcomings
  + The system was developed over time with no coherent design principles
  + Initially built to calculate index values but extended to include price collection and validation
  + Resulted in complex and unaudited discussion
* What did we want from a new system?
  + Flexibility to support different types of data (survey, administrative)
  + Adaptive data validation
    - Previous tools did not allow for any adaptation
  + Differentiation between price collection and data processing
    - Different disciplines which require different systems
  + Able to perform chain-linking
* Steps on the process
  + Assess
    - Documented and understood existing system and functionality
    - Done by business analysts
    - To find new methods and not replicate existing flaws
  + Requirements
    - Scoped out requirements and methodology for new system
    - “Blueprints” for all system build and testing work
  + Suitability
    - Researched suitable platforms in which to build system
    - Built data collection and data processing in separate environments to best suit the requirements of each discipline
* Development
  + Build
    - Code for the new system based on requirements using test data
  + Test
    - User testing performed by a dedicated team
    - Purpose to create test scenarios to ensure all requirements were met
    - Documented in full to ensure confident delivery to the end user
  + Engage
    - With stakeholders and end users throughout the development
    - Ensure that all user requirements were met and communicated
* Bringing it all together
  + Delivery
    - New system ready for use in production of Business Price statistics
    - New users supported and provided with guidance notes that facilitated the integration of the new system within their monthly publication cycle
* Ensure that the system was set for the future and could handle better analysis and new practices

*Discussion*

* Very good presentation. What have been the productivity gains in your production process?
  + Huge amount of productivity benefit. Less time taken away from the team in needing to deal with manual quality adjustments.
  + Almost able to entirely shift production team from doing checks to doing analysis – led to large gains in responding to the data and providing data to the end data users.

# Day 4 – September 22, 2022

**Cross cutting topic (7) – Data gaps and measurement issues with the online intermediary platforms (Air BnB, Uber, etc.)**

**Session leader: Ulla Virtanen (Finland)**

* Perspectives of the presentations
  + Both describe prices related to online platforms
  + The services under question differ
  + Switzerland – price of the intermediated service – user of the services (customer)
  + Singapore – price of the intermediary service provided – service or provider (customer)
* New initiatives
  + Challenges on data acquisition from global online platforms – EU perspective
    - No existing EU legislation granting access to transaction data of platform operations
    - Individual countries experiencing difficulties on establishing data exchange
    - Solution: Eurostat helping get access to the data
      * Tweet March 5, 2020
  + Agreement between platforms and Eurostat
    - Regular and reliable data from the platforms
      * Aggregated data shared with NSIs
    - Privacy guaranteed for guests and hosts
    - Publication of data
  + Why the motivation to share data?
    - Free publicity
    - Higher profile when done with Eurostat
    - Appearance of cooperative company
* Data contents
  + Breakdown
    - Geographical
    - Mode of accommodation
    - Size of accommodation
    - Residence of customer
  + Variables (quantities)
    - Number of stays
    - Length of stay
    - Nights spent
* National usages
  + Publish new experimental statistics (on more granular level than Eurostat)
  + Supplement survey data for tourism statistics (to include holiday rentals)
  + Lack of price data limits the usage
  + Drawback
    - No prices available as considered too sensitive

**A contrast between indices produced with Booking.com data and direct survey data. Swiss case study for the regions Bern and Lucerne SPPI. - Angela Hernandez Santacoloma (Switzerland)**

* Presentation Landscape
  + The hurdle – a gap between an index being produced from different sources
  + The Wind up – a new hotel industry index & |Covid-19
  + The situation in facts – with the use of the collected data and literature
  + The path ahead
* A NEW HOTEL INDUSTRY INDEX
  + Create a robust hotel index
  + Booking.com total 27 hotels. 12 Bern, and 15 Luzern
  + Survey total 29 hotels. 13 Bern, and 16 Luzern.
* COVID-19
  + 7.9 million in overnight stays for the big Swiss cities.
  + A reduction in overnight stays of 65% among the major Swiss cities Geneva, Zürich, and Basel.
  + -15.8 million in overnight stays for the Swiss hotel industry.
* The Hotel Reservation System
  + Direct Booking
    - Hotel reception, telephone, walk-ins
    - Traditional travel agency
  + Internet and e-commerce provides new platforms
    - Social media
    - Aggregators search engines
    - OTA’s
* Two elements to consider
  + The purpose of the hotel is to set the price close to the maximum willingness to pay of its customer’s, for which the hotel takes into account various factors available in real-time and then set the final price at the point of reservation. All the process takes place in real-time through digital systems.
  + Factors influencing hotel room reservation:
    - Date of arrival
    - Time of booking
    - Advanced booking (beforehand between six months and a week before arrival)
    - Availability and demand for a type of room
    - Hotel location
    - Length of stay
    - Number of hotel stars
    - Quality of the hotel (review)
    - Channel used to book the room (OTA’s, online aggregator, hotel website, distribution system).
* The Observations
  + Offline and online product prices are behaving differently, even though it is the same products
    - “Personalized discount”
  + Survey data and booking prices behave very differently
  + 2014-2022
    - Cyclitic up until 2020
* Most Important take away
  + It is probable for some branches that Internet price and in-store prices do not behave the same
  + When working with Internet data, consider if all prices behave closely:
    - Online list prices
    - Online transaction prices
    - In-store list prices
    - In-store transaction prices
  + Preliminary solutions:
    - At the end of the year evaluation will be done to see about publishing the data
    - First, a simple transaction prices index is collected directly from the hotels in a traditional manner.
    - Second, a hybrid form of a composed index including both transaction and online listing prices indices since they are both valid

**Singapore’s Experience in Developing Online Marketplaces SPPI - Edwin Boey (Singapore)**

* Introduction
  + The Singapore Department of Statistics (DOS) has been compiling the Computer Consultancy and Information Services Price Index (CISPI) since 2016. In the rebasing of the CISPI from base year 2016 to base year 2021, DOS included online marketplace services into the index structure, after incorporating classification changes in the Singapore Standard Industrial Classification 2020.
  + Online marketplaces are intermediary platforms that facilitate orders for goods or services between buyers and sellers usually on a fee or commission basis, without taking ownership of the goods or providing the services that are being sold through their online platforms (e.g., website, mobile application).
  + Some of the commonly used terms to describe intermediation services:
    - Virtual marketplace • Online marketplace • Multi-sided marketplace (or platform) • Two-sided marketplace (or platform) • Electronic markets (used in NAICS, Australia and New Zealand industry classifications)
  + DOS uses the term “Online Marketplace” instead of “Intermediary Platforms” as it is:
    - More commonly used in local media
    - More generally understood to referring to a website or application that facilitates transactions from different sources
* Classification
  + The Singapore Standard Industrial Classification 2020 (SSIC 2020) adopts the basic framework of the International Standard Industrial Classification of All Economic Activities Revision 4 (ISIC Rev.4) for international comparability, with appropriate modifications and updates to account for changes in Singapore’s economy and users’ requirements.
  + In the SSIC 2020, a new 3-digit group 632 ‘Online Marketplaces’ was created under Section J ‘Information and Communications’, with new five-digit Sub-classes
    - SSIC 2018 63: Information Service Activities 🡪 SSIC 2020 63: Information Service Activities and Online Marketplaces
    - SSIC 2020 6320 – Online marketplaces
      * For goods (including food)
      * For health services
      * For education services
      * For travel services
      * For real estate services
      * For services not else classified
  + Given the growing importance of different forms of online marketplaces and in the absence of international guidance on classifying such intermediaries at the time of our review in 2019-2020, DOS explored several options during the stakeholder consultations with data source and policy agencies before deciding to create a new code for ‘Online Marketplaces’ under Division 63 in Section J ‘Information and Communications’. Other options explored included:
    - Classify all online marketplace activities under SSIC 6312 ‘Web Portals’ in Section J.
    - Classify to existing/ appropriate SSIC codes based on the industry the online marketplaces are supporting (e.g., wholesale, retail) and create new codes under each SSIC section, where appropriate for all other intermediation services
* Market Conditions
  + Information Services account for around 6% of total revenue in 2020
  + Information services grew in both revenue and value added between 2019 and 2020
  + Distribution of revenue within online marketplaces for 2019 and 2020
* Development of SPPI for Online Marketplaces
  + 2020 – plan resources, timeline for developing SPPI, Internet research, Industry survey
  + 2021 – preliminary price surveys, clarify and refine survey items, determine pricing methods
  + 2022 – finalise firms, service products and weights. Compile 2021 Q1 to 2022 Q1, routine data collection and compilation.
* Survey of data collection
  + Asked for items such as monthly subscription, product listing/publishing fees, unit rates, other service items
* Possible Data Gaps and Measurement Issues
  + Definition/ Classification of Online Marketplaces
    - Determining if the services provided by the respondents are online marketplaces (e.g., application stores, dating applications)
      * More research and discussion undertaken by the team before confirming if it’s in-scope
      * Dating applications do not offer dating services but provide access to the application
    - Wrongly classified establishments that should belong to other SSICs (e.g., internet search engines, travel agents)
      * Feedback internally within DOS to review if these establishments should be classified elsewhere
    - Establishments that had online marketplaces but were classified in other SSIC codes.
      * Not included in the coverage but feedback provided internally within DOS to explore if there was a need to create a special division in SSIC 632
    - Larger establishments may derive revenue from non-online marketplace activities, such as sale of own goods/services via their platforms.
      * Not considered materially impactful and weights assigned to products based on total revenue
  + Data Collection / Compilation
    - For commission model, DOS is unable to efficiently utilise the survey returns provided by respondents by seller as the fees received vary drastically from month to month because of changes in products sold.
      * Could originally not be used to look at unit prices
      * Had to switch to generic commission rates by product categories or unit value (for all clients).
* Conclusion
  + DOS will monitor the existing products and prices, while looking towards engaging other online marketplace establishments to understand their service products and onboard them on the price surveys. More service products and prices are needed over a longer period to improve the methodology/ quality of this index and compile sub-indices by 5 digit SSIC.
  + Need more time to bring more firms on board and continue expanding the methodology

**Discussions on Cross cutting topic (7) – Data gaps and measurement issues with the online intermediary platforms (Air BnB, Uber, etc.)**

**Presentation from Task Force on criteria for fitness for use of alternative data - Martin J. Beaulieu (Canada), Scott Kilbey (UK)**

* Guidelines for Incorporating alternative data sources in official statistics
  + Determined the challenges faced by NSIs in assessing the quality of new data sources
  + Aligned alternative data literature and practical experiences
  + Developed practical tool
* Key considerations
  + Alternative data does not equal administrative data
    - Ex. Web-scraping is not administrative data
  + External data vs. Internal data
  + Structured vs. Unstructured data
  + High control vs. low control
  + Data ethics
    - Who is approaching who – is the data provider approaching the NSI or the opposite?
    - What is the impact?
  + All of the above consideration should influence how an NSI incorporates a new data source into their statistical processes. It is important to acknowledge these differences and to tailor the practical tools round the data source
* Variety of data sets and perceived risks
  + Countries should consider the array of possible data sources and the perceived risks that accompany them
  + Look at regulation, control, structured, consistency, metadata
* Generic Statistical Business Process Model
  + Version 5.0 of the GSBPM was updated “to be less survey-centric and [include] activities related to working with non-statistical data providers”
  + It is not a rigid framework in which all steps must be followed in a strict order
    - Is it brought in to fill in a gap or to expand coverage?
* Fitness for use questionnaire
  + Help NSIs determine the impact of adding new dataset
  + Help to ensure quality of adding in new dataset
  + Steps to integrate new data source
  + A series of questions in draft for to help NSIs navigate alternative data sources
  + Collection – how to collect the data
  + Process – are standard concepts and/or classifications being used in the data file? If not, how will this be addressed?
    - Will it be linkable with already processed datasets?
* Discussion and further work
  + This task force recommends that
    - This discussion be added to the agenda for Voorburg 2023
    - Member countries experiment with the proposed questionnaire approach using existing and new alternative data sources in order to provide feedback on its utility
  + Feedback from member countries
    - The tool developed by the task force leaned on
      * GSBPM
      * Statistics Canada quality Guidelines
      * Guide to reporting on Administrative data quality (New Zealand)
      * Internal documentation from Statistics Canada
      * The experience of task force members

**Discussion on criteria for fitness for use of alternative data**

* How many presenters are you looking for in 2023?
  + The more the merrier
* Is the risk of interruption of data availability included in the "control" category?
  + Longevity of the data source is of consideration when looking at new data sources
  + Want to ensure it is a data source that you can get over time
  + Something to consider adding to the paper before it is finalized
* Is increasing costs of purchasing data sets included in the considerations?
  + Similar to the data ethics – is the supplier approaching the NSIs and could then raise the price later on?
  + Questionnaire: Is there a cost to the data? Is there a justified need for the cost?
* Is there any way to approach large organizations who could give advice on this topic?
  + Not knowing what has been done to the data is a challenge to using alternative data. Potentially inheriting sources of error from the supplier. Should be tied into the acquisition process to help monitor the process.
  + Large businesses do not usually want to give away their secrets. Could possibly be approached but it would be very difficult.

**Cross cutting topic (3) – Trade asymmetries and imports of services John Jeremy (UK)**

* What are trade asymmetries?
  + Trade asymmetries: difference between estimates of partner countries when recording a trade flow
  + Exports recorded by one country and the imports recorded by partner country should be identical
    - Rarely the case – this is the asymmetry
  + Trade asymmetries are inevitable
    - Consequence of a range of conceptual and measurement differences
    - Becomes a problem when they are very large
  + Example; USA estimates credits (exports) to UK of 20bn, UK estimates debits (imports) from USA of 16bn
    - UK import asymmetry (UK perspective) is 4bn
* Current ONS approach to Asymmetries
  + Understanding Service Types
    - Current focus on Imports of Services
    - How to businesses behave?
    - How is service provided changing?
      * Are we keeping up with this?
  + Understand methods/data differences
    - What is the impact of these differences?
    - Looking at home country methods/data to other countries
    - What does that help us understand about our asymmetries?
  + Bringing Statistical Compilers together
    - OECD Expert Group on Asymmetries
      * Aim of group is to tackle asymmetries in a practical way – create achievable outputs
    - Important to have insight from the other countries
* Evolving Evidence/Knowledge base
  + Outputs from ONS published on ONS website
    - Available to international community
  + Deliverables from OECD group being developed
    - Options being considered
    - Considering developing a Wiki
    - Looking to feed into future compilers guides
  + Similar research to deflator development
    - Opportunities to share work
    - Promote consistency
* How can we help each other?
  + Sharing research – avoid duplication
  + Keeping everyone updated on pertinent issues
    - Changes to services
      * Highlight interesting or important changes
      * Perhaps informal linkage between OECD and Voorburg group
    - New data sources
      * How to do on international scale in order to minimize asymmetries
  + Consider wider SNA compilation
    - Consistency between areas of National Accounts
    - Consistency in trade statistics compilation
* Actions Proposed
  + Sharing information as development progresses (where possible)
    - Provide access to ONS/OECD/Voorburg findings, minutes, etc.
  + ONS provides update at Voorburg meetings
  + Considering an informal link between the OECD Expert Group and Voorburg Group
    - OECD Group is still taking shape
  + Emphasize the importance of considering wider statistical outputs
    - Members to actively consider and promote this

*Discussion*

* Is it possible that two NSIs could use the same data sets for their trade statistics?
  + Focus is not on removing the asymmetries, but on both countries estimating their data the best way that they can.
  + USA and Canada have done this and got their asymmetries down to 0
  + Manipulating the data down to 0 asymmetries is not necessarily the end goal as it may tell you something when you have the difference between the two countries
* How many sources do you have in your country? In Peru, there is few information and not a proper breakdown.
  + Struggle with import data. Easier to work with export data looking at what companies are producing in the UK but imports are more difficult because anyone can import products.
  + A lot of data comes in through surveys covering most of the service sector.

**Session: VG 2023 meeting planning**

* Virtual Voorburg 2022 Meeting Evaluation
  + All delegates will be emailed a link to a secure website after the meeting from VirtualVoorburg2022@statcan.gc.ca
  + Complete the short survey within one week of the meeting
  + Mexico will compile the results and they will be used to make improvements to future meetings
* Voorburg Group Bureau - End of Voorburg Group meeting 2022
  + We would like to thank Ramon Bravo from INEGI (Mexico) for his many years of service on the Voorburg Bureau.
  + We would also like to thank Dragos Ifrim from Statistics Canada for his great work as Voorburg Secretary.
  + Bonnie Murphy and Marcus Friden will remain as a co-chairs.
  + Welcome new Voorburg Bureau members
    - Mathieu Thomassin volunteered to be the secretary
    - 2 Vacancies with two nominations confirmed and vote confirmed they will be added
* Proposed Agenda 2023
  + Country Progress Reports/Alternative data industry spreadsheet
    - Very important and should be filled out by all even if you do not use alternative data
    - Send to Bonnie, Marcus, Mathieu, or Rohan
  + Data Collection topic: Quarterly measure of operating income, expenses and profit from combining a new survey of 400 big enterprises with administrative data from the Swedish tax Authorities
  + Chile and Mexico Collaboration on Financial Services (if ready)
  + Report from the new 2023 SPPI/CPI Task Force on comments from the presentation to the UNECE
  + National Accounts topic: Update on improvements to National Accounts – update and lessons learned from changes to the National Accounts processing form the UK deflator gateway that is scheduled to be implemented in 2023
  + Issues paper on Education
  + Collaborative guidance paper on SPPI and output compilation during economic shocks
* Industry based sessions – participants needed
  + Digital services – discuss issues included in Marshall Reindorf recent paper about digital economy inflation (weight update timing with respect to the pandemic, quality adjustment associated with transition to digital services, etc.)
  + Mini-presentations on industries in sector R (excluding division 90) – Arts, Entertainment and Recreation
* Methodology based sessions – participants needed
  + Summary of changes coming out of the System of National Accounts (SNA) and Balance of Payments Manual (BPM) updates and how they might affect services price and output statistics
  + Country examples of how and how often to reoptimize/update samples for PPI including initial recruitment of reporting companies
  + Update of the SPPI Manual – examples: recent work regarding CPI usage (no longer a special case). There are other examples of progress that has been made that would warrant an update.
  + Task force – practical experience with the use of CPI. Supply use table analysis included to determine correlation. Complex case studies. Culminates in a presentation to UNECE.
  + Production of a deflated index of turnover (converting nominal revenues to real?)
* Data collection – participants needed
  + New methods and challenges employed by NSOs post-pandemic to collect data “electronically” in lieu of in-person. Example: combination of electronic questionnaires (online), email, CATI, CAPI, etc.
  + Practical experience using the checklist developed for the 2022 VG meeting for evaluating alternative data fitness for use
* Collaboration – participants needed
  + Newsletter team – rotating members produce 2 newsletters between meetings (January and May)
* Cross Cutting Sessions – participants needed
  + Task force focused on improving and scheduling collaboration sessions within the Voorburg group (key component of our strategic plan)
  + Machine learning/AI driven processes
    - Experiences of machine learning, AI driven imputation and quality adjustment processes. This may also include initial efforts to put together data and metadata for eventual model training.
  + Would countries appreciate having access to desensitised data training sets? Potential longer-term benefits when comparing data from country to country and for harmonising and/or making improvements in imputation and quality adjustment methods on a per industry/product basis.
  + Innovative methods for capturing quality change
  + Glossary update
* Import topic sessions – participants needed
  + Use of PPI time series of countries from which we import products/services as a proxy
  + Measurement challenges for import services
* Contributions for 2023 Meeting
  + Use the chat to volunteer
  + Remain on the Teams Call during the break to speak with the Voorburg co-chairs
  + Send an email to the Voorburg co-chairs by October 7 with your name, NSO and topic
    - Can volunteer now and confirm participation by October 7
* VG Bureau and Delegate Tasks Before Next Meeting
  + VG 2022 Follow Up Tasks
    - Co-chairs report to UNSC to be delivered by November 30
    - First draft of meeting notes from Secretary due December 1, 2022
      * Session leaders review and comment (if needed) by December 31, 2022.
      * Final notes to the co-chairs by January 31, 2023.
    - Updates to 2022 papers are due no later than January 15, 2022
    - Update glossary – with new terms introduced at this meeting (based on meeting notes)
    - Web/communications – determine communications platforms/tools that will be used for task force and collaborative project work between meetings
      * Decision will be made if Yammer will be used
    - Future agenda – continue to look for the most pressing issues for participants to maintain the relevance and usefulness of the work of the Group
* VG Host 2023
  + Ruth Vizner – Israel
  + Israel hosting for the first time
    - Located along the western shore of the Mediterranean Sea
    - Mediterranean weather
    - Can travel the country in one day
  + Tel Aviv
    - Non-Stop City
    - Perfect weather year-round
    - 3 universities, 14 research hospitals
    - Start-up city – world’s leading tech hubs
    - Many popular visitor experiences including a great variety of restaurants
* Thank you
  + Thank you to Statistics Canada for hosting a very successful virtual meeting this year
  + Thank you to all VG participants for your many contributions and excellent insights, discussions and participation