

RESEARCH METHODOLOGY NOTE

Research Methods for the CREW Country Diagnostic Reports

August 2013

SUBMITTED BY



NATHAN
ASSOCIATES INC.

**Nathan Economic Consulting
India Private Limited**

SUBMITTED TO



Consumer Unity & Trust Society

Contents

1. Introduction	2
2. Overall Research Approach	4
3. Determining the Market Structure	6
4. Identifying Competition Enhancing Reforms	6
5. Research Methods	6
A. Defining the Market	6
B. Identification of Stakeholders	7
C. Determination of Hypotheses	8
D. Identification of Variables and Data to be collected	8
E. Data Collection	11
F. Analysis	14
Appendix A: Sample Case Studies	20

Tables

Table 1: <i>Staple Food</i>	7
Table 2: <i>Bus Transport</i>	7
Table 3: <i>Potential Data Parameters to Measure Consumer and Producer Benefits/Welfare</i>	9
Table 4: <i>List of Indicators</i>	10
Table 5: <i>Example Calculation Sample Size</i>	13

1. Introduction

The project entitled *Competition Reforms in Key Markets for Enhancing Social and Economic Welfare in Developing Countries (CREW)* is being implemented by CUTS International (CUTS) with research support by Nathan Economic Consulting India Private Ltd (Nathan India) over the course of three years starting in 2013. Supported by DFID (UK), BMZ (Germany) and facilitated by GIZ (Germany), the aim of the CREW project is to develop an approach which can be applied in developing countries to assess the benefits of competition reforms on consumers and producers in specific product markets.

After an initial analysis of potential countries and sectors for CREW's focus, four countries and two sectors have been selected for analysis: the **passenger transport** and **staple food** markets in **Ghana, India, the Philippines, and Zambia**. In the staple food sector, CREW will examine the following product markets: rice in the Philippines, maize in Zambia, maize in Ghana, and wheat in India. Similarly in the passenger transport sector, CREW will focus on bus transport for both inter-city and intra-city routes in the four countries.

The CREW analysis is expected to be conducted in three phases.

- The first is a *diagnostic phase*, which includes research on the institutional environment, reform efforts, and market structure of these sectors in the four countries and an assessment of the impact of reform on consumers and producers.
- The second phase will be the *design phase*, in which a “toolkit” will be developed to analyse the impact of competition reform in these markets across all four countries.
- The final phase will be the *validation phase*, in which the methodology will be applied and adjusted to finalise a framework for analysis of the impact of competition reform in developing countries.

This paper is part of Phase I, the diagnostic phase, and describes the research approach that will be utilised for preparing the Diagnostic Country Reports (DCRs). The objective of the DCRs is to examine the regulatory, institutional, and market environment of the chosen industries within the two sectors and conduct an initial examination of the impact of competition enhancing reforms on consumer and producer groups.

Since the DCRs will be prepared by different groups of researchers in the four CREW countries, this paper presents the general research approach and suggested methods to be applied across the countries and sectors. Specific research approaches will depend on the data available in each country and the research hypothesis being tested given the regulatory and market structure in each country/sector. Data constraints will mean that not all suggested methods are feasible. Hence, researchers of the DCRs will need to apply discretion in choosing the appropriate techniques, while following the general approach outlined here.

The aim of this paper is to provide a general and consistent approach with the aim of producing a set of DCRs from each country that are comparable, even if the specific research techniques applied may be slightly different. On the basis of the general research approach presented in this paper, the country researchers/partners will develop country-specific DCR research methodology notes, which will be

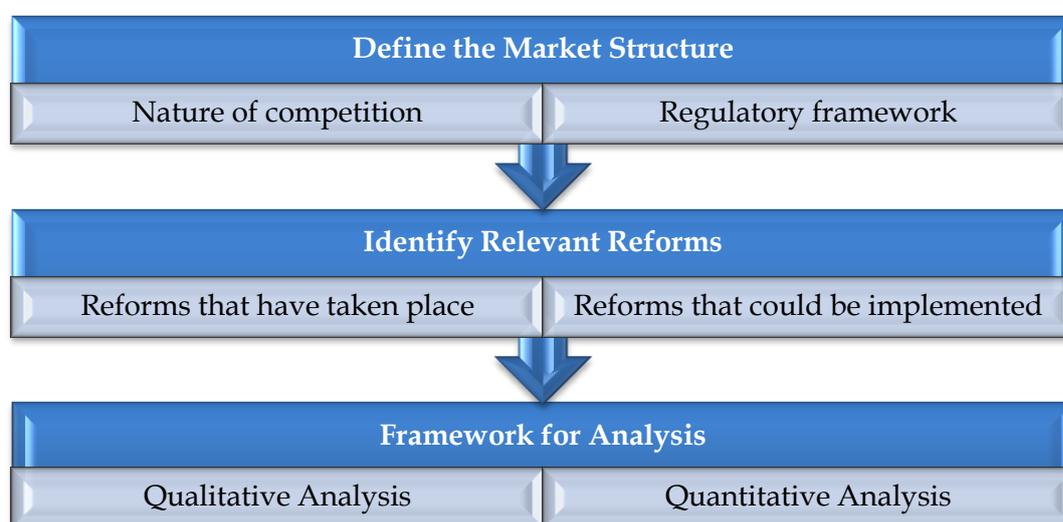
reviewed by CUTS and finalised by the in-country teams before the actual work on developing the country-specific DCR can begin.

This paper supplements other CREW documents to help researchers apply consistent research techniques to examine similar research questions across the product markets in all four CREW countries. This paper describes the broad research approach envisioned for producing the DCRs and details the methods described in the CREW “Outline of the Diagnostic Country Report”, which provides further information about each section of analysis. Readers of this paper should also read the “Outline of the Diagnostic Country Report” as well as the CREW Background Paper entitled “Measuring Impact of Competition Reforms in Developing Countries: A Survey of Possible Approaches and Selection of Countries and Sectors for the CREW Project” (www.cuts-ccier.org/CREW/pdf/Background-Paper.pdf).

2. Overall Research Approach

The Nathan team recognises that effects of competition may be hard to measure, and even harder to attribute to specific policy reforms. Disentangling the effect of other reforms, economy wide trends, and other economic events that affect the performance of the industry is a challenge. Further, given that the impact of reforms may not be felt immediately, there is an issue with time-lines for assessing impact as well. Some of the impacts may be easily and appropriately measured by numerical indicators. Others are better captured through observation and discussion, and some numerical indicators cannot be interpreted without accompanying qualitative information. For this reason, the research design should use a ‘mixed-methods approach’ including both quantitative and qualitative methods.

Broadly categorised, the objective of each DCR will be threefold.



- At the first level it will provide a detailed background of the market structure, including the legal and regulatory framework. This will include examining the number and nature of players in the market, their bargaining power relative to each other and input suppliers, potential of new entrants, substitute possibilities. This level of analysis will focus on the prevailing state of competition in the two markets in the four CREW countries.
- Next, the DCR will identify the set of reforms that have already taken place in the market (using a certain reference year), and select those reform measures that would be considered for closer examination to assess benefits for consumers and producers. Further, existing concerns of consumers and producers would also help tease out future/prospective reform possibilities and their benefits.
- Finally, after identifying the reforms, the DCR will provide the analysis framework to more clearly define the market of interest, identify players and stakeholders who will/could be affected (consumers and producers), frame the hypotheses for testing and define the counterfactual, identify the data required, data sources, and methods to test the hypotheses. The purpose of this analytical framework would be to assess benefits of past/ongoing reforms; and/or estimate those

from future/prospective reform actions. Where reforms have already taken place, ex-post analysis will be undertaken; where reforms have yet to occur, an ex-ante analysis will look at what would happen in the sector if reforms were to be undertaken. In some cases, a combination of ex-post and ex-ante analysis may be warranted.

The DCR will examine the following basic questions:

- What is the nature of competition in the market being analysed? Who are the producers? Who are the consumers? What structural forces are in play that promote or impede competition in the market?
- What legal, regulatory, institutional, or other reforms are in play, or could be in play, that impacts the competitive landscape in this market?
- How do the reform measures affect the market structure or competition in general? What impact (positive or negative) do these measures have on producers and consumers, if any?
- What are the quantifiable benefits and costs of these measures? Are benefits pro-poor?

The primary framework for analysis to answer questions about the competitive landscape of the market will be informed by the DFID Competition Assessment Framework. This framework has been widely cited to help identify key impediments to competition. Consistent with this framework, the DCRs will utilise “Porter’s Five Forces of Competition” as a frame of reference during the data collection and analysis phases to examine the intensity of competition in each market.¹ This model builds on the idea that competition in a market is not only determined by the extent of competition amongst current competitors, but also by the potential for new entrants into the market, the effect of substitute products, bargaining power of consumers, as well as bargaining power of firms that supply to the market.

In assessing the market structure, the DCRs will analyse the supply chain for each product/service, including pre-production, production, processing and distribution of a good/service. Within this framework, the DCRs will not only examine the market structure, but also legal and regulatory provisions that enable or discourage competition at each stage of the chain. If there are overall or sector-specific reforms that have been undertaken, the DCRs will elaborate on the content, timeline, and relevant stakeholders of such reform initiatives. If no reforms have taken place, the DCRs will identify needed priority reform/s based on stakeholder consultation and likely implications of such a reform/s on the market using ex-ante analysis. Practices at every node of the supply chain would also be analysed to assess their implications from a customer-supplier perspective, and also to trace impact on the ultimate consumers and producers as well.

With a comprehensive picture of the competitive and legal/regulatory landscape, the DCRs will then analyse the impact of current or possible reform efforts on the market, and consequently the effect on consumers and producers. Specifically, the research will investigate the impact on consumers by using such indicators as prices (including those of substitute products or services), product quality, range of choice available, change in coverage of product or service, and any non-price savings (e.g. in terms of time, effort, etc.). The impact on producers will be assessed in terms of change in innovation that

¹Five forces of competition’ model described by Michael Porter in *Competitive Strategy*, 1980

affect cost either directly or through suppliers; productivity gains; number of suppliers; number of direct competitors or producers of substitute goods/services; ease of entry/exit; ease of doing business; changes in profit; and any other relevant indicators.

Simply stated, an impact assessment is the comparison of key parameters or performance indicators before and after the intervention, for instance, in the context of CREW, the implication before and after the reforms in the particular sector. Several techniques, both quantitative and qualitative, can be used to assess impact of interventions. The appropriateness of the techniques will depend to a great extent on the availability of data and the characteristics of the parameters or performance indicators being evaluated.

An impact assessment requires estimating “the extent to which changes in the well-being of the target population can be attributed to a particular programme or policy”.² It compares changes in key indicators as observed with the programme to an estimate of the changes that most likely would have occurred anyway (the “counterfactual”).³ Impact may not always be positive; some reforms intending to enhance competitive may have actually had the opposite effect.

As stated in the CREW Background Paper, this analysis will utilise a mix of quantitative and qualitative research techniques. At a minimum, the DCRs will investigate secondary data that may be obtained from government agencies, such as governmental statistical offices, relevant line ministries, business associations, donor agencies, or independent think tanks. Data may also be collected through primary research, such as focus group discussions, perception surveys, and other anecdotal reports to fill key informational gaps not sufficiently addressed by existing information, or to supplement or build case studies around research that is already available.

The next three sections describe the objectives of the DCR: defining the market structure, identifying competition-enhancing reforms, and providing framework of research methods that may be used for analysis.

² Centre for Global Development, 2006, 12

³ USAID *Programming for Growth Briefing Note 8: Economic Impact*, June 2010, produced by Nathan Associates Inc.

3. Determining the Market Structure

The first step of the DCR will be to determine the market structure for both of the relevant markets – bus transport and staple foods – and the current regulatory structure governing the markets. Understanding the market structure is an important first step as it will allow the research teams see where competition enhancing reforms have occurred, or could occur.

Market structures can take various forms. Some of the most common market structures include:

- *Perfect competition*: A market with many buyers and sellers where both producers and consumers are price takers.
- *Oligopoly*: A market with a small number of interdependent sellers. Oligopolies can include oligopolistic competition with a few firms fiercely competing for consumers or cartels where firms collude and act in essence as a monopoly.
- *Dominant firm with competitive fringe*: One large, dominant firm⁴ with a few small competitors that can be considered as either an oligopoly with firms of unequal sizes or monopoly with new entrants.
- *Monopoly*: A market with only one producer and no close substitutes so that the producing firm has market power is able to set prices above marginal cost and earn a monopoly profit.

Each of the above market structures brings up different competitive issues. For example, a cartel may be operating in an oligopolistic market, restricting production acting as a monopoly, or a monopoly may be a regulated monopoly where the government or regulator sets prices.

The *DFID Competition Assessment Framework* will be the basis for assessing the competitive landscape in the chosen sectors in each of the four countries. The DFID Competition Assessment Framework focuses on:

- Defining the relevant market
- Identifying the markets and competitors
- Examining the market structure
- Looking for barriers to entry (or exit)
- Looking for anti-competitive conduct
- Considering vested interests and the principal beneficiaries
- Identifying government policies or institutions that limit competition

In conjunction with the DFID Competition Assessment Framework, we will use “*Porter’s Five Forces of Competition*” to recommend industries’ strengths and weaknesses:

1. Nature of competition in the industry
2. Potential of new entrants into industry

⁴ The OECD defines a dominant firm as one with at least 40 percent market share. See <http://stats.oecd.org/glossary/detail.asp?ID=3199>.

3. Power of suppliers

4. Power of customers

5. Threat of substitute products

In order to identify the government policies or institutions that limit competition, the researcher must first identify which policies and regulations are governing the market. This will require desk research and interviews with stakeholders. The possibility for applying other relevant tools will also be examined.

4. Identifying Competition Enhancing Reforms

Next, the DCR will identify the competition-enhancing reforms that have already taken place, or could take place, which would promote competition. The DCR will identify elements of competition reforms to be examined in each sector starting with identifying reforms which are targeted or likely to have an impact on competition. On the basis of literature review and discussions with CREW project advisers, 'competition reforms' for the purposes of this project has been defined as an aggregate of:

- Enabling government policies that promote competition in markets.
- Appropriate regulatory framework and institutions for promoting competition in sectors including sector-specific regulations.

Effective competition enforcement institutions. In choosing reforms to evaluate, it is important to think about the underlying policy objective of the reform and what the reform is trying to address. The OECD's Competition Assessment Toolkit⁵ can be used in choosing reforms by looking at whether the reforms/recommended reforms:

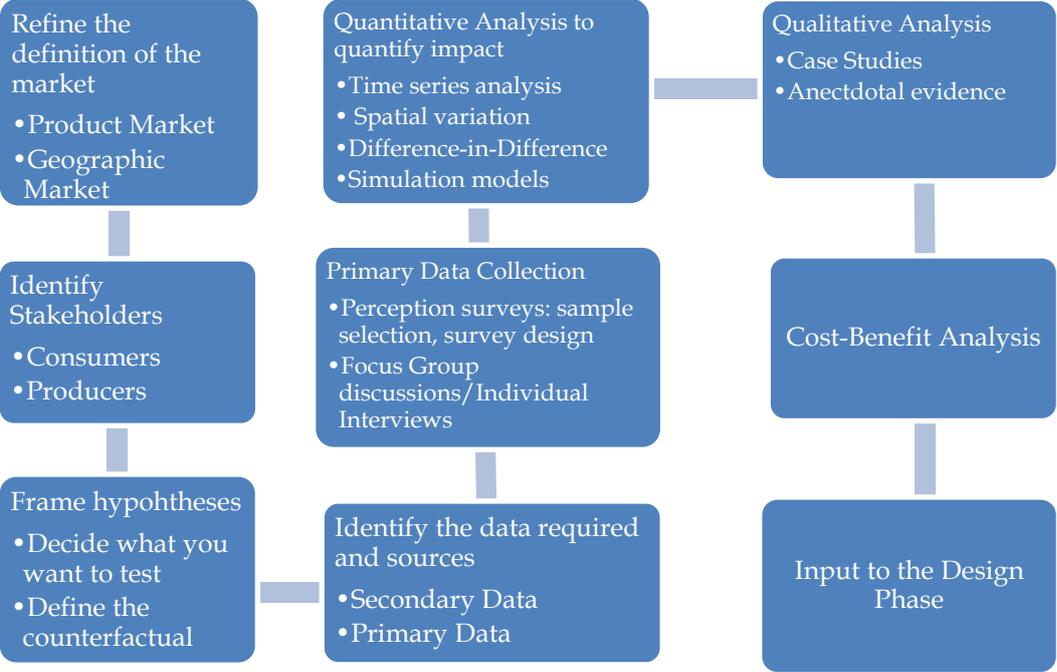
- Remove barriers to entry and increase number and range of suppliers
- Allow suppliers to compete
- Incentivise suppliers to compete
- Increase investment

In selecting reforms to study, we will have to be mindful of the time period of the reforms. The focus should be on reforms that have taken place in the last 10-15 years, keeping in mind that more recent reforms might not yet have produced their full impact, in which case both ex-post and ex-ante analysis may be combined to estimate both impact that has occurred, and impact that will continue to take place. In some cases the researcher may not find any substantial reforms that have taken place to date; in this scenario, the researcher should identify a specific reform that could benefit competition if it were implemented, and conduct ex-ante analysis. The next chapter of this paper will detail various research methods that are applicable.

⁵<http://www.oecd.org/daf/competition/46193173.pdf>

5. Research Methods

After identifying the reforms, the DCR will provide the analysis framework to define the market, identify players and stakeholders who will/could be affected including both consumers and producers, frame the hypotheses for testing and define the counterfactual, identify the data required, determine data collection methods including data sources, sample selection and survey design, and methods to test the hypotheses. These steps are further described below.



A. Defining the Market

The market of interest has to be clearly defined so that impact analysis can be focused. A market definition includes both the definition of the product market and geographic market. While in a broad sense the markets for assessment have already been chosen to be a specific staple food and passenger bus transport, a further refinement is necessary at this stage. It is unlikely that data will be available to do a quantitative analysis, and such a rigorous definition is not necessary for the purpose of our reports, but a qualitative assessment should be conducted so that the analysis is both complete and focused.

The product market should be determined in the view of the products at hand – whether it be rice, wheat, maize or bus transport. A review of substitute products should be considered just in the sense to determine whether potential or actual substitute products compete with the products at hand, constraining the behaviour of producers. To do so, the products would need to be at a similar price point and quality such that consumers would switch to these substitutes if prices of the product of study substantially increased. For example, in the Philippines passenger bus transport market, the DCR should consider whether jeepneys provide a competitive constraint to bus transport operators.

The geographic market must also be defined in the context of our study and the chosen reforms. For example, are we looking at a state, city or national level reform or regulation? In the case of bus transport, the relevant geographic market of interest should also specify the routes of study. In staple food, the DCR should discuss whether it is looking at the whole country or specific states/cities as a market.

B. Identification of Stakeholders

Once the market for study is clearly defined, the next step in the DCR process should be the identification of relevant stakeholders who either affect or are affected by the reform. On the consumer side, this study will focus on the end-consumer, but on the producer side, will look at all levels of the supply chain from pre-production to distribution. Producers also include both incumbents and new entrants to the market, and these two stakeholders might face different scenarios in terms of changes in their welfare. While stakeholders in each sector or product market may vary depending on the country, some potential stakeholders may include:

Table 1: Staple Food

Consumers	Producers	Other
General public (preferably differentiated by income group)	Pre-production (such as banks/MFIs lending to farmers) and input suppliers (such as fertiliser or seed suppliers)	Think Tanks/University Research Centres
	Farmers	Government regulatory bodies
	Millers	Farmers/Business Associations
	Storage facilities	Consumer Advocacy Groups
	Distributors (Public or private)	

Table 2: Bus Transport

Consumers	Producers	Other
General Public	Bus operators (Public and Private)	Regulatory Bodies
	Input suppliers (e.g. Tires, petrol/diesel)	Bus terminal operators
		Operators' Association
1.	2.	Think Tanks/University Research Centers
		Consumer Advocacy Groups

Relevant stakeholders are then mapped along a supply chain of the product market being considered, to clearly illustrate the particular stage of the pre-production, production, or distribution stages where the impact of a particular reform measure is likely to be felt.

C. Determination of Hypotheses

Once a clear understanding of the regulatory and competitive landscape emerges for a product market in a country, a pertinent hypothesis should be developed which asks to what extent a reform measure affected (or could affect) competition in the market, and how it affected (or could affect) consumers and producers. In statistics, setting up and testing a hypothesis is an essential beginning step and allows the research team to understand which variables should be collected for analysis.

In the case of CREW, we can test various hypotheses regarding the effects of competitive reforms on various measures of competition in terms of consumer and producer welfare to arrive at our goal of this project, which is to better demonstrate measurable benefits from effective competition reforms in developing countries for ensuring long-term support for competition. As part of determining the hypotheses for the DCR, it is necessary to clearly articulate the reform whose impact is being assessed, either focusing on one relevant reform measure or a concise set of reforms, depending on which is more feasible to measure and likely to have produced (or may produce) a significant and measurable impact.

For example, consider that a reform is implemented in an economy where the government was previously providing goods, now allowing private players to enter the market. In order to test the effect of the reform, the null hypothesis would be “the effect of entry of a new player does not have an impact on the price of the goods provided” and the alternative hypothesis in this case could be “the effect of entry by a new player has an impact on the price of the goods provided”. This hypothesis can be tested using various statistical techniques commonly used for hypothesis tests – parametric and non-parametric—with variables such as the number of players, supply and demand factors, prices and an indicator variable for the reform.

While defining the hypotheses, it is also important to use qualitative analysis to define the counterfactual. That is, what would have occurred absent the reform?

D. Identification of Variables and Data to be Collected

Once a researcher has determined what they are looking to achieve and measure through their research, the next steps are to determine what data are needed to make these measurements, and how to achieve collecting the necessary data. In the real world, data availability is never as straightforward as the researcher would hope in an ideal world, and compromises have to be made to account for realities of time, money and data availability. Data can be collected from a combination of primary (survey, focus groups) and secondary (census, government records, international statistics) sources depending on availability.

In the case of CREW, the relevant hypotheses should be measuring effects of competition on consumer welfare, and benefits to producers/producer welfare on the basis of the following parameters listed in Table 3.

Table 3: Potential Data Parameters to Measure Consumer and Producer Benefits/Welfare

Consumer Welfare	Producer Benefits/Welfare
Access: Have the reach of goods and services to consumers in areas where they were not available earlier (distributive efficiency)?	Enabling Environment: Is there some predictability of regulatory actions, so that prices may be adjusted if feasible?
Quality/Price: Are better quality goods and services available for consumers? Has consumer safety improved due to better quality product? Have prices increased appreciably of such good/services?	Productivity: Have new innovations, such as improved ICT tools or dedicated bus lanes, led to increased productivity? Have increased productivity, reductions in paperwork, or decreased input costs led to cost savings?
Choice: Are new products available to consumers? Are new firms entering the market offering more variety?	Profits: Have reforms benefited producers by leading to higher producer/market profits?
Price: Have prices changed, either increased or decreased?	
Time savings: Do offered goods and services help save time for consumers?	

For example, increased transparency in the market or predictability of regulatory actions could reduce the cost of doing business therefore affecting producer welfare. These variables are just to give an indication of potential ways that consumers and producers can possibly be affected by actual or potential competition enhancing reforms.

We understand that some of the indicators may affect both consumers and producers, depending on who is considered a producer and who is considered a consumer. For this reason, we define consumers to only include end-consumers, and producers to include producers at all levels of the supply chain from pre-production to distribution. We also recognise that not all impacts on producers will be positive and reforms may lead to a reduction of producer welfare as welfare is transferred from producers to consumers or incumbents to new entrants. The parameters presented in Table 3 for Producer Welfare will be difficult to gather quantitatively. Therefore, they will rely more heavily on qualitative data collection. Finally, many other factors, such as access to essential services, fair market processes, etc. will likely affect most producers negatively.

Table 3 tries to identify those indicators (prices, costs, profits) whose effect on producer welfare is generally not ambiguous. Data on other supply-side indicators capturing the general business environment (such as improving access to essential services and increased transparency in the market) and the competitive framework (such as fair market processes/access) should be considered, but not necessarily quantified in producer welfare, as the results will likely be negative leading to interpretation problems if these factors are included in a producer welfare calculation.

In examining the welfare effects using the above parameters, the analysis will need to be mindful of net welfare effects,⁶ i.e. a reform measure which might have produced a welfare gain for consumers might have reduced the welfare for producers. That is to say, the welfare gain might just be a transfer of welfare between producers and consumers, rather than a net decrease in the deadweight loss. Similarly, producer welfare may be transferred from incumbents to new entrants, leading to higher welfare for some producers but lower welfare for others, or lower welfare for producers overall as increased competition leads to lower prices transferring welfare from producers to consumers.

Additionally, welfare effects in the short-run as well as the long-run must also be considered. For instance, a decrease in market prices in the short-run may produce a welfare gain for consumers in the short-run. However, if firms do not then earn the resources to invest in research and development, in the long run, this might result in a welfare loss for consumers because of lack of updated products in the market.

Nathan India recognises that data to quantitatively assess the trade-offs between short-and long-run welfare effects as well as net welfare effects between consumers and producers may not be easily accessible in the four identified CREW countries. However, this kind of analysis can be performed at very minimum using qualitative analysis techniques.

In order to analyse the above parameters and the hypothesis framed, suggested data to be collected through primary/secondary data collection include but are not limited to:

Table 4: List of Indicators

Staple Food	Bus Transport
Price data (domestic and import)	Price data by route (per km of travel, if available)
Costs of production (fertiliser, seed, land costs, labour costs etc.)	Costs of production (vehicle costs, driver wages, fuel costs, maintenance costs, tolls etc)
Quantities produced, quantities imported/exported, units of goods sold	Frequency and number of buses per route per day or week, length of bus routes
Number of distribution channels	Number of passengers per route
Number of types of products available at various levels of quality	Available roads and infrastructure
Availability of substitute products including prices and quantities of goods sold	Substitute products and their details including price, travel time, frequency, quality
Number of players in the market and their market shares for different levels of the supply chain	Number of players in the market (public and private) per route and their market shares
Changes in health outcomes such as child malnutrition statistics to measure changes in availability of food/affordability/quality	Number of fatalities and injuries over time due to bus accidents to measure changes in quality
Reports of consumer safety violations or illness/death due to poor quality/contaminated product/foodborne illnesses to measure quality	Passengers per bus/capacity of bus to measure overcrowding/quality
Any taxes, tariffs, or registration/licensing fees	Any taxes, tariffs, or registration/licensing fees

⁶ While net welfare effects may consider a transfer of producer welfare to consumer welfare as being neutral/having no net effect, we will view such an increase in consumer welfare as having a net positive effect. For example, the UK Office of Fair Trading’s “A guide to OFT’s Impact Estimation methods” only considers consumer welfare when evaluating impact. While we will also look at impacts on/benefits to producers, we will focus on consumers when looking at welfare.

To the extent possible, data collected should span a long enough timeframe to cover periods before and after (at least five years each) any reform measure which is being evaluated. If no reform have yet been implemented, data for a period of the last 10-15 years will be preferable.

E. Data Collection

3. SOURCES

In general, the data can be collected from secondary and primary sources. In this regard, as a first step, data should be collected from secondary sources. An illustrative but non-exhaustive list of possible sources of information includes:

- Competition Authority, if existent
- Ministry of Commerce/Trade
- Government censuses
- Business Associations
- International Donor Agencies such as the World Bank
- Local research institutes
- Consumer advocacy groups
- Publicly available online sources

In scanning available data from secondary sources for indicators that measure producer and consumer welfare as outline in Table 3 above, a researcher may find that data key to answering the hypothesis question may be missing or faulty. In such cases, a researcher may need to conduct primary data collection, such as surveys or focus group interactions. The scope of such a data collection depends on a number of different factors, including the variables to be measured, budget or other resource constraint, usefulness of data so collected to the overall research question, etc.

4. SAMPLING METHODS

When collecting primary data via surveys or focus group discussions (FGDs), a sampling strategy must be determined indicating who will be surveyed and how the respondents will be chosen. It is difficult to create a “one size fits all” sampling strategy since the best strategy is often dependent on the characteristics of the population in the individual country/product market being surveyed. In the case of CREW project, a single methodology should be employed in each country to reduce biases in the cross-country comparison. Some basic sampling strategies are described below.

Probability sampling involves the random selection of sample elements from the sampling frame and such a sample can be used to make inferences about the target population using statistical methods. However, probability sampling requires that the elements in the sample have known probabilities. Types of probability sampling include simple random sampling and stratified random sampling. When simple random sampling is used, every frame element has an equal probability of inclusion, giving unbiased estimates without weighting. If random sampling is used but the probabilities vary, weighting can be used to make adjustments. When there are distinct subgroups within the population with great variation across subgroups, it is important to use stratification instead of random sampling to ensure that all subgroups are sampled and represented in the data. Stratification divides elements

into population subgroups, ensuring representation from each, which will lead to less variance of estimates. Below are the basic steps undertaken in stratified sampling:

- Take sampling frame of agent list
- Divide frame into mutually exclusive groups (strata)
- Independently select the sample agents from each strata

Strata can be proportionate or disproportionate to that group's percentage of the population; disproportionate strata are used when the sampling rate varies across strata to allow for adequate sample sizes for independent analysis of subgroups (oversampling). When using disproportionate strata, one must post-weight the data to adjust it to be representative of the population before conducting the analysis. Using either method produces data that can be used to make inferences about the population.

Clustering is used when populations are spread across large geographic areas, which may be too costly to sample randomly. A key consideration is that there may be intra-cluster homogeneity due to the environment, self-selection and interaction between elements within that cluster. Steps undertaken in clustering include:

- Divide frame into clusters
- Randomly choose clusters for your sample
- Within the cluster, choose randomly agents to be in the sample to survey

While stratification usually lowers variance, clustering increases variance but lowers costs. The two can be used in combination. For the CREW project, Nathan India recommends a broad strategy which includes a stratified random sampling method,⁷ with clustering to control costs if the geographic areas are large. For staple foods, strata could be defined based on region and urban/rural. For transport, since a few major routes will be studied, we recommend covering both urban and rural areas. While a nationally representative sample is usually preferred by researchers, the cost of collecting such a sample must be considered and weighed with the benefit of doing so. In the case of a large, diverse country like India, a nationally representative sample is cost-prohibitive, and as such, the focus should be on just two or three states.

If probability sampling is not possible, there are other potential methods to use, but these may limit the ability to do statistical analysis as the data sample might not be representative of the population which means that you cannot use the sample to infer about the population. Despite these limitations, non-probability sampling may be the most feasible option in some cases, especially in the case of bus transport. Examples of non-probability sampling are accidental sampling (such as 'man on the street'), snowballing, and purposive sampling.

⁷ A stratified random sample is a sampling plan in which a population is divided into mutually exclusive and exhaustive strata, and a simple random sample of units is taken within each stratum. The sampling is performed independently within each stratum.

Source: Stratification and Stratified Random Sampling, Georgia Southern University

5. SAMPLE SIZE

A sample size should be large enough to ensure statistically significant results while maintaining cost-effectiveness. As each CREW country has different characteristics, the ideal sample size and specific sampling strategy will be determined by the local survey agency. In general, the formula to calculate sample size is:

$$n = \frac{\left[\frac{P(1-P)}{\frac{A^2}{Z^2} + \frac{P(1-P)}{N}} \right]}{R}$$

where,

- n is the sample size,
- Z is the z-value at the desired confidence level (1.96 for 95 percent),
- P is the estimated variance of an attribute that is present in the population (in case of a large population where we do not know the variability in the proportion for practical reasons p is assumed to be 0.5),
- A is the precision level i.e., the margin of error allowed (say 5 percent),
- R is the response rate (in general 70 percent or 0.7), and
- N is the total number of population.

Below is an indicative table showing the number of sample according to the population as per the above formula:

Table 5: Example sample size calculation

	Example 1	Example 2
Total population of people in the sampling frame (N)	100,000	500
Degree of variance (P)	0.5	0.5
Precision desired (A)	0.05	0.05
Confidence interval (Z)	1.96	1.96
Response rate (R)	0.7	0.7
Sample size (n)	547	310

6. SURVEY DESIGN

The primary data for the analysis will be collected through perception surveys and/or FGDs. The CREW team will design questionnaires for both producers and consumers in each of the sectors. While we recognise that slight modifications to the questionnaires will have to be made on a country-specific basis, the questions in general should be similar across countries. To ensure consistency, CUTS will be providing the basis of the questionnaire to the implementing agencies in all four countries, which then can be modified on a country specific basis and translated to local languages as required. The survey canvassing would be approached via a combination of email, phone and in person interviews depending on the on-the-ground logistics and capabilities of the implementing partners.

Survey data can be collected using either both pen and paper interviewing (PAPI) or computer assisted personal interviewing (CAPI). CAPI can include either computers or Personal Digital Assistant (PDA) devices or smartphones, which are often a better option in developing country context as they are more mobiles than laptops and are easier to use in remote locations. We recommend using CAPI when possible as it reduces data entry errors and cost as the data is entered on the spot, which reduces the level of effort of the survey team. This also reduces errors as the validation occurs while the data is being entered, which helps enumerators in filling out the correct option and following skip patterns (which also cuts down on interview time). Real-time analysis can also be completed to check the quality of the answers received. CAPI saves on printing, also improving logistics. The pitfalls of CAPI may include upfront costs for the hardware and software, training costs to ensure that the enumerators are familiar with the devices, and potential misplacement of the devices. Also, steps must be taken to ensure the data is not lost due to technical issues or power outages and that the enumerators have a reliable source of power to keep the PDAs sufficiently charged.

Finally, both the enumerators and the population being surveyed must be familiar enough with the devices so that there is no mistrust of them, which has been an issue with household surveys in India. In cases where CAPI is not possible, PAPI can be used, but will require a data entry tool for entering and validating the survey data. The tool should limit the responses such that only responses appropriate to the question can be filled in and should also ensure that all questions are answered. PAPI requires more skilled enumerators, detailed training, and oversight in order to prevent data collection errors.

In either case, the survey team must undergo sufficient training prior to the survey. The training should include a run-through of the questionnaire, including detailed explanations of all of the questions and potential answers, and mock interviews. Sufficient training ensures that the surveyors are comfortable with the questionnaire, which will decrease the risk of interviewer bias, interview time, and measurement error.

The field work should begin with a few days of pilot surveys, which test the questionnaire in the local language. After the pilot field work, a few days should be allocated to analyse the results, make changes to the questionnaire, re-print the questionnaire or re-program the PDAs, and train/inform the surveyors of the changes. After the pilot has validated the questionnaires and survey team, the field work will take place. Spot checks and follow up calls should be done to check a percentage of the field work. Finally, it is important that surveyors keep track of the people that they attempted to survey, in order to allow the data analysts to calculate the response rate.

F. Analysis

We will use a variety of techniques - a mixed methods approach - to measure impact including multiple regression techniques, time-series analysis, spatial variation, simulation models, case studies, anecdotal evidence from stakeholders and cost-benefit analysis. While the exact methods used to measure each hypothesis will be determined on a case-by-case basis, it is important that the same

techniques are used to test each hypothesis across countries, so that a cross-country comparison is possible.

1. QUANTITATIVE ANALYSIS

Quantitative methods that we recommend using for evaluating the benefits of the reforms on competition can be broadly categorised in the following way.⁸

1. Time series variation - Comparing outcomes before and after a reform
2. Spatial variation - Comparing outcomes between reformed and unreformed markets, or markets with two different reforms
3. Difference-in-difference-Comparing outcomes between two markets over two periods of time
4. Structural estimation/ simulation models – estimating the potential effects of a reform
5. Cost-benefit analyses – Comparing monetised costs and benefits of a reform

The first three approaches deal with economic impact analysis, where a scenario with regulation is compared with what might have happened without the regulatory constraint (counterfactual). Simulation models can be used to estimate the impact of a potential reform in markets where reforms have just or not yet occurred. On the other hand, cost-benefit analysis (CBA) is commonly used to assess the net benefits of a past or proposed reform by monetising all the relevant benefits and costs. CBA is often used in conjunction with other methods such as regression analysis to quantify and attribute benefits.

While using any of the assessment techniques, it will be important to distinguish impact of reform from other external shocks. Controlling for as many variables as possible which might impact the dependent variable in a regression analysis, and assigning appropriate and rational proportions of costs and benefits attributable to the reform in a CBA will be important. Assessment of the impact is also complicated by the significant time lag before impact is fully realised and the catalytic effects of interventions that yield benefits not clearly attributable to a programme. In most cases, assessment techniques depend on the availability of data and/or the feasibility of data collection.

a) *Time Series Analysis*

When time series analysis is used to compare two samples of data before and after a regulatory reform is introduced there is an assumption that, all else being equal, changes observed with this before and after approach are attributable to the regulation introduced. The time period studied must include period(s) when the regulation is introduced. Data for the period after regulation may be realised in reality (*ex-post*) or may be estimated based on what is expected to be realised (*ex-ante*). In the former case, ideally the data spans a long enough period to allow for transitional changes to be fully realised. In this approach a dependent variable is defined (e.g. price, rate of technical change, etc.) and modelled as a function of exogenous economic variables. A dummy variable⁹ is assigned for whether a data point is from the “regulated” or “unregulated” pool. The sign and coefficient on the dummy corresponds to the impact and size of the regulation.

⁸ See the CREW Background Paper for more detail.

⁹ This variable is constructed such that it only takes two values - zero or one, depending on whether the data point is from the unregulated or regulated pool respectively. Economists have traditionally called such a variable a “dummy” variable.

Using time series data before and after a reform to estimate the impact of competition policy reform is one of the simplest and expedient methods of analysis but one must keep in mind data availability limitations, carefully choose data time periods, and use econometric rigor to avoid estimation errors. Understanding the unique circumstances of the regulatory reform in question is crucial in order to correctly place the pre and post data points and correctly assess impact of reform. A more detailed discussion on the potential pitfalls of time series data is discussed in the discussion paper.

b) Spatial Variation Analysis

Similar to the time-series approach, a spatial variation method compares two sets of markets which are identical or similar in all respects, except for the regulatory constraint introduced in one. These pairs could be markets in different states or countries, for instance, or customers impacted by a particular regulation and those that are not.

c) Difference-in-difference

Difference-in-difference enriches spatial variation techniques by incorporating time series elements along with spatial variation to compare changes in outcomes from two groups over time. The term refers to measuring the impact of reform as a difference in outcomes from the outcome realised without one. The method looks at the change from “before” to “after” across two groups, which typically includes a treatment group and a control group. As mention above, the two groups could be markets in different states or countries where one did not have a reform and the other did. The reform and control pools should have comparable trends before the reform.

Cross-sectional data requires both a defined variation in the sample groups, as well as control for relevant non-regulatory variations. As with the time-series variation, a dummy is assigned for the “regulated” and “non-regulated” pools and regressed as a function of the exogenous explanatory variables.

While this method produces more precise estimates of the effect of policy reform than basic time-series regression, this method is quite data intensive. It is often difficult to find two sample sets that are characteristically similar, except for the regulatory reform being studied. Again, this is further discussed in the discussion paper.

In the context of the CREW project, difference-in-difference is most applicable and feasible to conduct in India, where there are a multitude of states with different policies which provides a good setting for a natural experiment. One way to undertake such an analysis would be to compare two states, one where a competition enhancing reform has taken place (treatment group), and the other where no such reforms have taken place (control group). The states should be as similar as possible in terms of culture, economy, demographics, size etc. other than the reform. Alternatively, instead of having a treatment and control states, two treatment states that underwent different competition enhancing reforms (but started with similar situations) could also be compared. Another interesting possibility would be to look at two different states that underwent similar competition enhancing reforms but had different outcomes, to see which external political economy, cultural or social factors made the reform less productive in one of the states. Before choosing which difference-in-difference approach to use, the researchers must first conduct secondary research on approximately 10 states to determine for

each sector 1) the competition-enhancing reforms that have taken place in the last 10-15 years in that state, and 2) the demographic, economic, cultural and social factors that are present in that state.

In all of these three methods, an issue faced while measuring impact is controlling for external factors, which is done by using multiple regression techniques whenever data allow. These regression techniques might include a mixture of ordinary least squares (OLS), two-stage least squares (2SLS), difference-in-difference, and limited dependent variable techniques such as logit or probit models. Regression analysis allows for the analyst to hold all factors equal between interventions to determine whether the intervention had a positive or negative statistically significant effect on the dependent variables. One of the biggest challenges of this task will be controlling for and separating out the impact of different interventions that took place at the same time.

In cases where data are not available to conduct regression analysis, we will use non-parametric tests to calculate frequencies and compare the mean before and after the reform to measure the impact. While these methods do not allow us to control for other factors that could have influenced outcomes, they are still a good way to estimate impact in the face of data constraints.

d) Structural estimation/ simulation models

An alternative approach uses economic models that can run simulations of equilibrium for counterfactuals to determine the effect of regulation. Structural econometric models detail the economic and statistical assumptions required to estimate economic quantities modelling either the entire economy (general equilibrium) or a portion thereof (partial equilibrium). In structural models, economic theory is used to develop mathematical relationships between a set of observable “endogenous” variables to another set of observable “explanatory” variables and unobservable variables, which may not always be estimable. Such simulations are most relevant to the CREW project in the context of ex-ante analysis that simulates what the effect would be if a reform were enacted.

e) Cost-Benefit Analysis (CBA)

In cases where we have found that reforms have already occurred, after quantifying the impacts (using the above techniques), we will determine which impacts are attributable to the various reforms and convert them to present value USD terms to undertake cost-benefit analysis. In cases where reform are expected, we can estimate what the likely benefits and costs would be based on answers to survey questions as well as other data from secondary sources. Cost-benefit analysis is used to measure the impact of an intervention in terms of the economic benefits it generates or gains to the economy, relative to the costs of the intervention or losses to the economy, both actual costs and opportunity costs (what could have been achieved with the capital and labour had they been used elsewhere) discounted over time.¹⁰ This will allow us to measure the cost of certain interventions compared to the value of the impact achieved to see if the impacts are in fact significant given the costs.

¹⁰ Brown, Richard and Campbell Harry, *Benefit-Cost Analysis: Financial and Economic Appraisal Using Spreadsheets*, Cambridge University Press, 2007: pp. 62-78.

The following mechanism would be employed in computing the CBA, depending upon the availability of data and time:

- Calculation of social costs, which includes the calculation of both the tangible and the intangible costs. These costs are not easily measured but can possibly be quantified in terms of costs of inefficient firms failing, reform implementation costs such as the cost of setting up commissions, subsidy costs or cost savings due to the removal of a subsidy, costs to producers such as costs due to increased safety regulations etc.
- Calculation of social benefits, which includes the calculation of both the tangible and the intangible benefits. These can be quantified in terms of increased tax revenues, reduced fatalities, injuries or illnesses, improvements in health outcomes, reductions in time etc.
- Converting all costs and benefits to USD terms.
- Discounting the future value of benefits, as the costs and benefits accrue over time.
- Comparison of the costs and benefits to determine the net social rate of return.
- When costs and benefits cannot be directly compared in terms of USD, we will calculate the cost per unit, i.e. the cost per additional bus on a route or cost per passenger.
- A qualitative approach would further be used whenever benefits would be intrinsic and not measurable in monetary terms such as improvements in quality, innovations, and new products.

As is common practice, we recommend expressing the results of the cost-benefit analysis in terms of the economic rate of return (ERR), the net present value (NPV), or the benefit-cost ratio. These results can be used to compare outcomes across countries and sectors.

Conducting cost-benefit analysis will add another dimension to the survey and secondary data collection component of the impact assessment, thus allowing for a multi-dimensional analysis of the reforms and basis for a cross-country comparison.

2. QUALITATIVE ANALYSIS

In the absence of sufficient data to undertake a purely quantitative evaluation of meaning, we will need to rely on qualitative information generated by surveys, interviews, or FGDs. These surveys or discussions seek to understand the value of the reforms for intended beneficiaries, which will be coded to be used in a quantitative framework or used to infer qualitative information about programme results. For example, data that measure quality of service of buses, or the quality of food grain, may not be easily available. We can use answers to survey questions about quality (e.g. by rating quality on a numerical scale of 1 to 5) for quantitative analysis. Similarly, we may not be able to quantitatively measure staple food shortages before vs. after a reform, but anecdotal evidence from people's real life experiences can provide evidence that food shortages are less (or more) common since the reform. These stories may complement more general quantitative data such as nutrition statistics to allow us to judge the impact of the reform despite the absence of a strict quantitative measure.

Qualitative analyses will also pose open-ended questions about reform strengths and shortfalls, which are often a source of new information on benefits or potential ideas for future reforms. These participatory methods will allow us to determine impact in areas such as changes in behaviour.

Information from these interviews or discussions will be corroborated against any available quantitative information to provide a full analysis.

FGDs include a discussion with a sample group in order to understand their in-depth experience (in this case, it would be the opinion about the reform and its implications in respective sector). The information gathered from FGDs provides insight explaining the results of the quantitative analysis, and can provide a basis for case studies.

A case study is a study which highlights the experience of an individual, group, or company, and clearly showcases the actual plot of the sample area. For example, in the case of staple food, a useful case study that each implementing partner should conduct would be to follow one particular supply chain from the farmer to consumer. Sample case studies are included in Appendix A.

We recommend organising qualitative analysis including FGD among the stakeholders at various stages. In this regard, separate questionnaires have to be designed for the FGD considering the variables where there is no quantitative measurement.

Appendix A: Sample Case Studies

CASE STUDY: FOLLOWING A STAPLE FOOD SUPPLY CHAIN

- I. Background
 - a. Specific geographic area of study
 - b. History of staple food supply in the area
- II. Market Structure
 - a. Structure of the market
 - b. Supply chain including flowchart and the competitors at each level
 - c. Time series of statistics including production quantities, trade statistics, consumption, prices, tariffs/taxes
- III. Competition Enhancing Reforms
 - a. What is the specific reform of interest and what level of the supply chain is it focused on?
 - b. What is the context in which the reform was passed – political scenario, policy goals?
- IV. Following the supply chain and assessing the situation and impacts at each level:
 - a. Pre-production: input suppliers (fertiliser, seeds) and financing/banking
 - b. Production, farmers
 - c. Millers
 - d. Storage
 - e. Distributors/marketers
 - f. End-consumers
- V. Conclusions and Inferences
 - a. Was the reform successful, if so, at what levels and why?
 - b. What were the changes in consumer welfare?
 - c. Did producer welfare increase as well? At what level?
 - d. What policy implications can be made?
 - e. An assessment of external validity – can the reform be applied in other areas or countries?
 - f. What challenges still exist? Are they being addressed?

CASE STUDY: BUS TRANSPORT MARKET STUDY

- I. Background
 - a. Specific route of study
 - b. History of transport in the area
- II. Market Structure
 - a. Structure of the market
 - b. Supply chain including flowchart and who are the competitors at each level/their market shares
 - c. Time series of statistics including number of passengers, frequency of buses, prices, costs, average travel time
 - d. Regulatory Framework
- III. Competition Enhancing Reforms
 - a. What is the specific reform of interest and what level of the supply chain was it focused on?
 - b. What is the context in which the reform was passed – political scenario, policy goals?
- IV. Following the supply chain and assessing the situation and impacts at each level:
 - a. Pre-production: input suppliers (vehicles, fuel), financing/banking, infrastructure, bus terminal operators
 - b. Bus companies
 - c. End-Consumers
- V. Conclusions and Inferences
 - a. Was the reform successful, if so, at what levels and why?
 - b. What were the changes in consumer welfare?
 - c. Did producer welfare increase as well? At what level?
 - d. What policy implications can be made?
 - e. An assessment of external validity – can the reform be applied in other areas or countries?
 - f. What challenges still exist? Are they being addressed?