

Contraceptive Forecasting Accuracy: Trends and Determinants

Ali Mehryar Karim
Karen Ampeh
Lois Todhunter



DELIVER
No Product? No Program. Logistics for Health

Contraceptive Forecasting Accuracy: Trends and Determinants

Ali Mehryar Karim
Karen Ampeh
Lois Todhunter



DELIVER
No Product? No Program. Logistics for Health

DELIVER

DELIVER, a six-year worldwide technical assistance support contract, is funded by the Commodities Security and Logistics Division (CSL) of the Office of Population and Reproductive Health (PRN) of the Bureau for Global Health (GH) of the U.S. Agency for International Development (USAID).

Implemented by John Snow, Inc. (JSI), (contract no. HRN-C-00-00-00010-00), and subcontractors (Manoff Group, Program for Appropriate Technology in Health [PATH], Social Sectors Development Strategies, Inc., and Synaxis, Inc.), DELIVER strengthens the supply chains of health and family planning programs in developing countries to ensure the availability of critical health products for customers. DELIVER also provides technical support to USAID's central contraceptive procurement and management, and analysis of USAID's central commodity management information system (NEWVERN).

This document does not necessarily represent the views or opinions of USAID. It may be reproduced if credit is given to John Snow, Inc./DELIVER.

Recommended Citation

Karim, Ali Mehryar, Karen Ampeh, and Lois Todhunter. 2004. *Contraceptive Forecasting Accuracy: Trends and Determinants*. Arlington, Va.: John Snow, Inc./DELIVER, for the U.S. Agency for International Development.

Abstract

This study examines the accuracy of the contraceptive forecasts contained in the Contraceptive Procurement Table (CPT) database, which is maintained by JSI/DELIVER to evaluate the quality of the forecasting process, to assess the utility of the forecasts in procurement planning, and to monitor and evaluate project progress. One-year-ahead contraceptive forecasts for 50 clients from 19 countries for the years 1995 to 2003 are analyzed. Among others, the results indicate that the forecast accuracy has been improving over time, which is partly attributable to the improvement of the client's logistics management information system and the use of the PipeLine software for preparing the CPTs.

DELIVER

John Snow, Inc.
1616 North Fort Myer Drive
11th Floor
Arlington, VA 22209 USA
Phone: 703-528-7474
Fax: 703-528-7480
Email: deliver_project@jsi.com
Internet: deliver.jsi.com



Contents

Executive Summary	vii
Introduction	vii
Principal Findings.....	vii
Discussion	ix
Chapter 1: Forecast Accuracy: Trends and Determinants	1
Introduction	1
Data and Methodology	1
Results	4
Discussion	24
Chapter 2: Influence of PipeLine and LMIS on Forecast Accuracy	27
Introduction	27
Analytic Framework.....	27
Data, Measurements, and Analytic Technique.....	29
Results	32
Discussion	36
References	37
Appendix	39

Figures

1. Frequency Distribution of the Percentage Difference between Projected and Actual Consumption of Contraceptives and the Distribution of the Forecast Error or Accuracy, 1995–2003 Pooled	3
2. Percentage Distribution of the Reported Quantity of Contraceptives Procured According to Donor, 1995–2003 Pooled	6
3. Scatter Plots between the Projected and Actual Consumption, 1995–2003 Pooled	6
4. Trend in the Mean and Median Forecast Accuracy for All Methods, 1995–2003.....	7
5. Trend the Median Forecast Accuracy by Method, 1995–2003.....	9
6. Median Forecast Error by Method, 1995 to 2003 Pooled.....	10
7. The Trend in the Median Forecast Error by Region, 1995–2003	11
8. Median Forecast Error by Client Category, 1995 to 2003 Pooled...	11
9. Trend in the Median Forecast Error for all Method by Client Category, 1995–2003	12

10. Median Forecast Error by Number of Donor, 1995 to 2003 Pooled	13
11. Trend in the Median Forecast Error for All Methods by Number of Donors, 1995–2003	13
12. Percentage of the Forecast That Overestimated or Underestimated Actual Consumption by More Than 25% by Background Characteristics, 1995–2003 Pooled.....	15
13. Trend in the Percentage Over- or Under-Forecasting, 1995–2003 ..	16
14. Trend in the Percentage Over- or Under-Forecasting by Client Category, 1995–2003	17
15. Trend in the Percentage Over- or Under-Forecasting by Region, 1995–2003	18
16. Trend in the Percentage Difference between the Aggregate Projected and the Aggregate Actual Consumption by Method, 1995–2003	19
17. Trend in the Percentage Difference between the Aggregate Projected and the Aggregate Actual Consumption by Region, 1995–2003	20
18. Trend in the Percentage Difference between the Aggregate Projected and the Aggregate Actual Consumption by Client Category, 1995–2003	21
19. Trend in the Percentage of the Proposed Shipment Realized in Actual Shipment by 75% or More in Quantity, 1995–2003	22
20. Percentage Under-, Average-, or Over-forecasting by Adequate and Inadequate Shipment, 1995–2003 Pooled.....	22
21. Trend in the Median Forecast Error by ‘CPT year’ (year 1), ‘CPT Planning Year’ (year 2), and the Year after ‘CPT Planning Year’ (year 3), 1995–2003	23
22. Scatter Plot between the Actual Consumption Reported within Two Subsequent CPT Years, All Methods, 1995–2003 Pooled	24
23. Trend in the Percentage of the Cases That Used PipeLine to Prepare the CPT, All Methods, 1995–2003	25
24. Logistics Cycle.....	27
25. Analytic Framework.....	28
26. Scatter Plots between Forecast Accuracy and LMIS Index Score, 1995, 1999, and 2000.....	34
27. Impact of PipeLine and LMIS on Forecast Accuracy	35
1A. Format of a CPT Report.....	39

Tables

1. Countries Included in the Forecast Accuracy Analysis	2
2. Description of the Sample.....	5
3. Median Forecast Error by Quintiles of the Quantity Projected, 1995–2003 Pooled	14
4. Forecast Accuracy at 25th, 50th, and 75th Percentile, All Methods, 2002 and 2003 (n=228)	26

5. Description of the Items Used to Construct the LMIS Capacity Index	31
6. Relationship between the Use of PipeLine and Forecast Accuracy, 1995–2003 Pooled	33
7. Description of the Forecast Error and LMIS Index Score by Year	34
1A. Trend in the Median Forecast Accuracy by Background Characteristics, 1995–2003 (sample size in parenthesis)	40
2A. Trend in the Mean Forecast Accuracy by Background Characteristics, 1995–2003 (sample size in parenthesis)	41
3A. Trend in the Median Forecast Error by Country, All Methods, 1995–2003 (sample size in parenthesis)	42
4A. Trend in the Percentage of the Forecast for All Methods That Overestimated or Underestimated Actual Consumption by More than 25% and the Percentage of the Forecast Within $\pm 25\%$ of the Actual Consumption, According to Background Characteristics, 19.....	43
5A. Trend in the Sum of Projected and Actual Consumption (in 1,000s) for All Countries Included in this Analysis and the Percentage Difference Between Them, by Background Characteristics, 1995–2003	44
6A. Regression Models Predicting the Effect of PipeLine on Forecast Accuracy (n=1,050)	45
7A. Comparison of the Characteristics of the Sample That was Selected for the Analysis of LMIS Effect on Forecast Accuracy and the Sample that was Not Selected (1995, 1999, and 2000 pooled)	46
8A. Regression Models Predicting the Effect of LMIS on Forecast Accuracy (n=207)	47

Executive Summary

Introduction

Contraceptive Forecasting Accuracy: Trends and Determinants examines the Contraceptive Procurement Tables (CPTs) in the NEWVERN database (maintained by John Snow, Inc./DELIVER) to determine the trend and determinants of forecasts accuracy. The analysis, discussed in this paper, evaluated the quality of the forecasting process, determined the utility of the forecasts in procurement planning, and monitored and evaluated the project's progress.

Methodology

Forecast accuracy (or error) is defined as the absolute percentage difference between projected and actual quantities of a contraceptive distributed in a specific year for a client or program. The analysis was limited to the CPTs for 50 clients from 19 countries (nine from the Africa, four from the Asia/Near East region [ANE], and six from the Latin America and the Caribbean [LAC] region) where JSI had considerable input in preparing the CPTs. One-year-ahead contraceptive forecasts for 1,050 CPTs prepared between 1994 and 2002 were validated for accuracy using the CPTs prepared between 1996 and 2004. Using regression methods, the statistical significance of the trend and the determinants of the variation in the mean and the median forecast accuracy were assessed. The analysis assessed the variation in the forecast accuracy by country, region, method, client category, number of donors, use of PipeLine software to prepare CPTs, and functional level of the CPT client's logistics management information system (LMIS).

Principal Findings

1. Results of the analysis show that between 1995 and 2003 the forecast accuracy has been improving at a steady rate. The median forecast error improved from 35 percent for the 1995 forecasts to 26 percent for the 2003 forecasts.
2. For contraceptive methods, the improving trend of the forecast accuracy was most prominent for the pill and the injectable.
3. Among the three regions, the declining trend in the forecast error was most rapid for the LAC region, followed by the Africa region. Although the declining trend in the forecast error for the ANE region was not very prominent, the overall forecast error in the ANE region was the lowest of the three regions.

4. Among the three client categories (i.e., Ministry of Health [MOH], social marketing [SM], and other NGOs), only the MOH clients showed a declining trend in the forecast error.
5. The trend of the forecast accuracy did not differ significantly when more than one donor supplied contraceptives compared to when only one donor supplied contraceptives.
6. After pooling the forecast accuracy measures over the analysis period, the forecast accuracy varied by method, region, client category, and projected quantity of a contraceptive.
 - Among contraceptive methods, the forecast error for implants was the highest; there were no significant variations in the forecast error between the condom, pill, injectable, and IUD.
 - Among regions, forecast accuracy was highest for the ANE region, followed by the LAC and Africa region.
 - For the client categories, the forecast accuracy was similar between SM and MOH clients; however, it was lower for other NGO clients.
 - The forecast error was greater when a smaller quantity of a contraceptive was being forecasted.
7. The contraceptive forecasts were more likely to overestimate rather than underestimate the actual consumption. The over- or under-forecasting varied by method, region, client category, and number of donors.
 - For contraceptive methods, over-forecasting was higher for IUDs when compared to the other methods, which were more or less the same.
 - For regions, over-forecasting was lower in the ANE region when compared to the Africa or the LAC region.
 - For client categories, over-forecasting was higher for other NGO clients compared to the MOH or the SM clients.
 - For the number of donors, over-forecasting was higher for single donors compared to multiple donors.
8. To assess the utility of the CPTs for global procurement planning by the U.S. Agency for International Development/Commodities Security and Logistics Division (USAID/CSL), the aggregate percentage difference between projected and actual use was assessed at the global level. The analysis indicated that there was a tendency to overestimate the actual consumption at the aggregate level by about 10 percent.
9. Projected shipment accuracy was also assessed. Shipment accuracy was defined in terms of its adequacy. If a client or program received 75 percent or more of the proposed quantity for shipment of a contraceptive, the client or program was considered to have adequate projected shipment accuracy. About two-thirds of the proposed contraceptives shipment of a contraceptive in the CPTs was adequately met. The projected shipment accuracy did not vary significantly over time nor did it vary by method, region, client category, or number of donors. The projected shipment accuracy was associated with forecast accuracy. Higher forecast accuracy was related to higher shipment accuracy.

10. Forecast accuracy was compared between the current year forecasts, one-year-ahead forecasts, and two-years-ahead forecasts. As expected, forecast accuracy was highest for the current year forecasts, followed by the one-year-ahead forecasts, and then by the two-years-ahead forecasts.
11. Implementing *PipeLine* software to prepare CPTs decreased the forecast error by an average of approximately six percentage points.
12. The association between forecast accuracy and the functional level of the CPT client's LMIS was assessed. The analysis showed that forecast accuracy improved when the LMIS performance of a client improved. A 42 percentage point reduction of the forecast error was attributable to the functioning level of the client's LMIS.

Discussion

The improved forecast accuracy could be explained by one or more of the following: (a) improvement in the forecasting methodologies and procedures followed by the logistics advisors, (b) improvement in the ability of the CPT clients to obtain historical dispense-to-user data due to an improved LMIS, and (c) increase in the frequency of forecasting. The analysis period for this study was the project life of the Family Planning Logistics Management III (FPLM III) plus the first two-and-a-half years of DELIVER. One noteworthy improvement in the forecasting methodology during this period was the introduction and use of the PipeLine software. In 1997, PipeLine was used for 13 percent of the forecast included in this analysis; the use gradually increased to 100 percent by 2003. PipeLine improved the efficiency of the CPT preparations through automation; it could have improved forecast accuracy by providing a way to consistently estimate the actual consumption of a contraceptive and by avoiding mathematical errors. The software also estimated the projected quantity of a contraceptive needed more frequently by the client. However, without an effective LMIS in each country/program to feed PipeLine the appropriate information, it is unlikely that the software alone could explain all the observed decline in the forecast error.

An earlier analysis by Gelfeld (2000) showed that the logistics management system of the countries where FPLM III provided technical assistance improved between 1995 and 2000, indicating that the declining forecast error during that period could be partly explained by the improved logistics system. Further analysis confirmed this hypothesis. Because FPLM III provided technical assistance to improve the logistics system capacity of the CPT clients, part of the impact of the LMIS on forecast accuracy could be attributed to the project's activities.

The recent forecast accuracy observed in this analysis sets a standard accuracy level for all CPTs in the future. The level of the median forecast error during 2002 and 2003 was 26.7 percent, with the 95 percent confidence interval between 23.1 and 30.5 percent. This is within the acceptable range (i.e., within 25 to 30 percent on average) for a one-year-ahead forecast, according to U.S. commercial forecasting standards.

It is a common practice for countries with multiple family planning programs to conduct forecasting and procurement plans separately, even for a contraceptive that is supplied by the same donor. Since lower accuracy is associated with forecasting smaller quantity of a contraceptive, splitting up the forecasting of the national requirement for a contraceptive (into smaller quantities) by different programs may introduce error. This may be avoided by conducting pooled forecasting for contraceptives for all programs in a country.

Although the aggregate forecast for a contraceptive in the CPTs slightly overestimates the aggregate consumption, by keeping the bias in mind, USAID/CSL can use the aggregate forecast to plan and monitor central-level procurement planning.

Chapter 1

Forecast Accuracy: Trends and Determinants

Introduction

Since 1986, John Snow, Inc. (JSI) has worked closely with the U.S. Agency for International Development Commodities Security and Logistics Division (USAID/CSL), and the Centers for Disease Control and Prevention (CDC) to prepare and maintain Contraceptive Procurement Tables (CPTs). The CPTs are used to reliably estimate contraceptive needs, develop appropriate procurement plans, and monitor contraceptive shipment orders for worldwide family planning programs that receive donor support (DELIVER 2000). The CPTs have been archived in a database, NEWVERN, which is currently being maintained by Central Contraceptive Procurement (CCP), DELIVER. This paper examines the CPTs to determine the trend and determinants of the variance between the projected and actual contraceptive use, i.e., the forecast accuracy. This analysis, the fifth of its kind, was conducted to evaluate the quality of the forecasting process and the utility of the forecasts in procurement planning, and to monitor and evaluate project progress.

Data and Methodology

CPTs are prepared annually to forecast or project the quantity of each contraceptive brand a program expects to dispense during the coming year. This information is used to propose procurement quantities and/or shipment orders for the required commodities. For this analysis *CPT planning year* is the year for which procurement plans are made, and the *CPT year* is the year before the CPT planning year. In addition to the contraceptive forecast for the CPT planning year, the CPT also contains the contraceptive forecast for the CPT year and for the year following the CPT planning year (see figure 1A). However, this analysis evaluates only the projected quantity of contraceptive use for the CPT planning year because of its greater interest to the client/programs and the contraceptive commodity donors for financing and procurement planning purposes. Contraceptive orders are usually not placed based on the forecasts for the CPT year or for the year following the CPT planning year.

The methodology for analyzing forecast accuracy was based on an earlier work by Wilson (1995). The level of forecast accuracy or error is measured by taking the absolute percentage difference between projected and actual quantities¹ of a contraceptive distributed in a specific year for a given client or program, using the following formula:

1. The actual consumption reported in the CPTs, on many occasions, was estimated by using the best available data (Wilson 1995). Therefore, it is possible that part of the forecast error is contributed by the error in estimating actual consumption.

$$\text{forecast accuracy} = \left| \frac{\text{actual consumption} - \text{projected use}}{\text{projected use}} \right| \times 100$$

The actual contraceptive use or consumption is defined as the quantity dispensed to the user during a specified period of time. For this study, the quantity of the actual consumption of a contraceptive is obtained from the CPT reported for the year preceding the CPT year. To obtain forecast accuracy, the forecasted quantity of a particular contraceptive use for the CPT planning year is compared with actual use, obtained from the two-years-later CPT report.

The analysis was limited to the forecasts made between 1995 and 2003 using the CPTs prepared between 1994 and 2002. Although about 3,170 CPTs from 59 countries were available in NEWVERN for the study period, the authors restricted the analysis to the 19 countries (see table 1) where DELIVER had considerable input in the CPT preparation processes. This ensured that the results of the analysis, to a certain extent, reflect the performance of the project. CPTs for vaginal foam tablets were excluded from the analysis because they are being phased out; CPTs for female condoms were excluded because there were very few CPTs for the relatively new product. Ultimately, the forecasts in 1,734 CPTs were used to conduct this analysis.

Table 1. Countries Included in the Forecast Accuracy Analysis

Country	CSL Priority	PEPFAR	Region	Client Category		
				MOH	SM	Other NGO
Bangladesh	X		APR		X	
Bolivia			LAC			X
Burkina Faso			AFR	X		
Cameroon			AFR	X	X	X
Egypt	X		ENE	X		
El Salvador	X		LAC	X		X
Ghana	X		AFR	X	X	X
Guatemala			LAC	X		X
Haiti		X	LAC	X		
Malawi			AFR	X		
Mali			AFR	X	X	X
Nepal	X		APR	X	X	
Nicaragua			LAC	X		X
Peru	X		LAC	X		X
Philippines	X		APR	X		
Tanzania	X	X	AFR	X	X	X
Togo			AFR	X	X	X
Uganda	X	X	AFR	X	X	
Zimbabwe			AFR	X		

Forecast accuracy for many of the CPTs was not estimated because it could not be linked with its two-years-later CPT. The reasons were because—

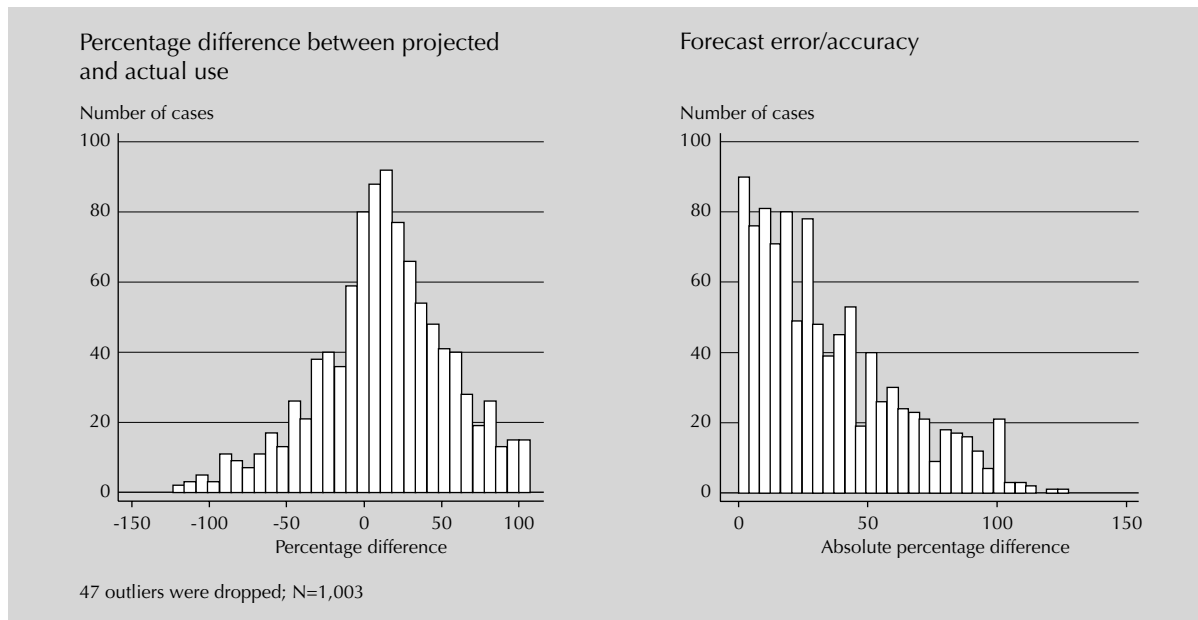
- the product or the program was phased out
- the program was no longer DELIVER's client
- the product brand name was changed between the two CPT years
- some of the 2004 CPTs were not done during the time of this analysis.

Matching the product type rather than the brand name between the two-years-apart CPTs prevented the CPTs/forecasts from being excluded from the analysis because the condom's brand name had changed. CPT forecasts were dropped if the projected use for a given product was zero, which made the forecast error undefined.

About 60 percent of the total 1,734 forecasts were validated, which gave a total of 1,050 cases of forecast accuracy or error measurements available for this analysis. The analysis represented 50 clients from 19 countries, including nine CSL priority countries (see table 1).

Descriptive statistics were carried out to describe the characteristics of the CPTs. Means and medians were used to describe the central tendency of the forecast error. However, the median was preferred over the mean as the summary statistics for describing forecast accuracy because of two factors: (1) the expected outliers and (2) by definition, the distribution of the forecast error or accuracy is skewed to the right. The distribution of the percentage difference between the projected and actual use² appears similar to a normal distribution (see figure 1). The transformation of the

Figure 1.
Frequency Distribution of the Percentage Difference between Projected and Actual Consumption of Contraceptives and the Distribution of the Forecast Error or Accuracy, 1995–2003 Pooled



2. Percentage difference between projected and actual use = $100 \times (\text{projected use} - \text{actual use}) / \text{projected use}$.

percentage difference between the projected and the actual use to the absolute percentage difference (i.e., the forecast error or accuracy) forced the negative errors from the left tail of the distribution of the former to the right tail of the distribution of the latter, which produced a right-sided skewed distribution of the forecast error.

The presence of outliers³ was expected in this analysis for several reasons that may or may not relate to the forecasting methodology, including (1) unforeseen inclusion or exclusion of another contraceptive product or service delivery system in the market that significantly decreased or increased the utilization of the reference product under study, leading to a larger than expected forecast error; (2) the methodology for estimating actual use may have changed over time—for example, the source of data for estimating the actual consumption may have changed over time—which could result in larger than expected forecast error; and (3) in some cases, the lack of adequate consumption data could have randomly introduced a larger-than-expected forecast error. These reasons, however, may not necessarily produce outliers, and could have contributed to the observed forecast error.

Median regression methods implemented by Stata (StataCorp 2003) were used to assess the trends and differentials in the forecast accuracy. All statistical tests were controlled for some of the other sources of variation in the forecast error. For example, the model that assessed the variation of the forecast accuracy over time was controlled for the variation of the forecast error due to variations in the countries, clients, and products between the forecast years. Further details on the models are provided as notes under the appropriate tables.

Results

Description of the Cases

Table 2 displays the characteristics of the forecast accuracy cases or the sample included in this analysis. The distribution of the cases/sample over the analysis period was similar for the forecast years 1995, 1996, 1997, 2000, 2001, and 2002 (about 12 percent on average); however, it was lower for the forecast years 1998, 1999, and 2003 (about 9 percent on average). Half the cases were from the Africa region (AFR), which represented 21 clients from nine countries; about two-fifth of the cases were from the Latin America and the Caribbean region (LAC), representing 24 clients from six countries; and the remainder (12 percent) were from Asia and the Near East region (ANE), representing five clients from four countries. About one-third of the sample was for oral pills; about one-fifth of the sample was for condoms, and one-fifth for injectables; 17 percent of the sample was for IUDs, and the rest (8 percent) of the sample was represented by implants.

3. Forecast errors greater than three standard deviation (1 Std. Dev. = 44.0) of the forecast errors for the last two years (i.e., 2002 and 2003) were labeled as outliers. Forty-seven outliers were identified.

Table 2. Description of the Sample

Characteristic	%	Characteristic	%
Forecast year		Client	
1995	12.8	Ministry of Health	60.4
1996	11.1	Social Marketing	13.7
1997	11.3	Other NGOs	25.9
1998	8.2		
1999	9.1	Single or multiple donor	
2000	12.0	Single	75.3
2001	13.8	Multiple	24.7
2002	13.2		
2003	8.5		
Region		USAID	80.1
Africa	49.8	UNFPA	26.8
Asia & the Near East	12.2	DFID	8.1
Latin America & the Caribbean	38.0	IPPF	5.9
Method		Sample size (n)	1,050
Condom	20.8		
Oral pill	34.3		
Injectable	19.8		
IUD	17.1		
Implant	8.1		

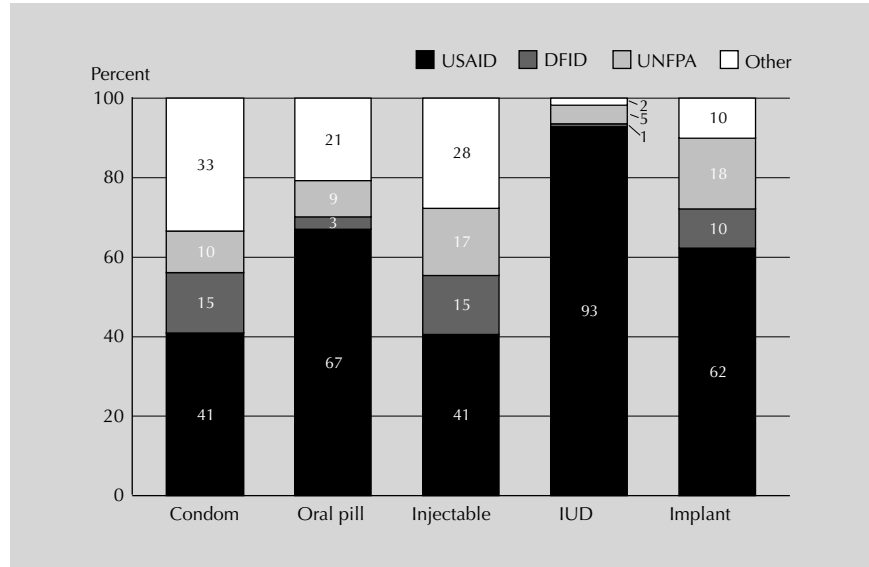
All 50 CPT clients included in the analysis were categorized into three groups: Ministry of Health (MOH), social marketing (SM), and other NGOs (including IPPF-supported NGOs). Most (60 percent) of the clients were MOH; about 14 percent of the sample represented SM clients, and the remaining quarter of the sample represented other NGOs.

The sample was also categorized based on the number of donors involved with procuring the forecasted contraceptive.⁴ A single donor was involved in most of the cases (75 percent). USAID supported the funding for contraceptives as a single donor or with others, in 80 percent of the CPTs; while UNFPA (27 percent), DFID (8 percent), and IPPF (6 percent) were the other major donors.

Figure 2 shows the percent distribution of contraceptives by source or donor pooled over the period 1995 to 2003. As expected, the major supplier of contraceptives that were procured based on the CPTs and included in this analysis was USAID, which provided almost all (93 percent) of the supplies for IUDs, and most of the supplies for oral pills (67 percent) and implants (62 percent). Approximately two-fifths of the condom and injectable supplies were supported by USAID. The next major donors for contraceptives were DFID and UNFPA.

4. The information on the donors was obtained from the shipment information contained in the two-years-later CPT that earlier provided the information on the actual consumption used for estimating forecast accuracy.

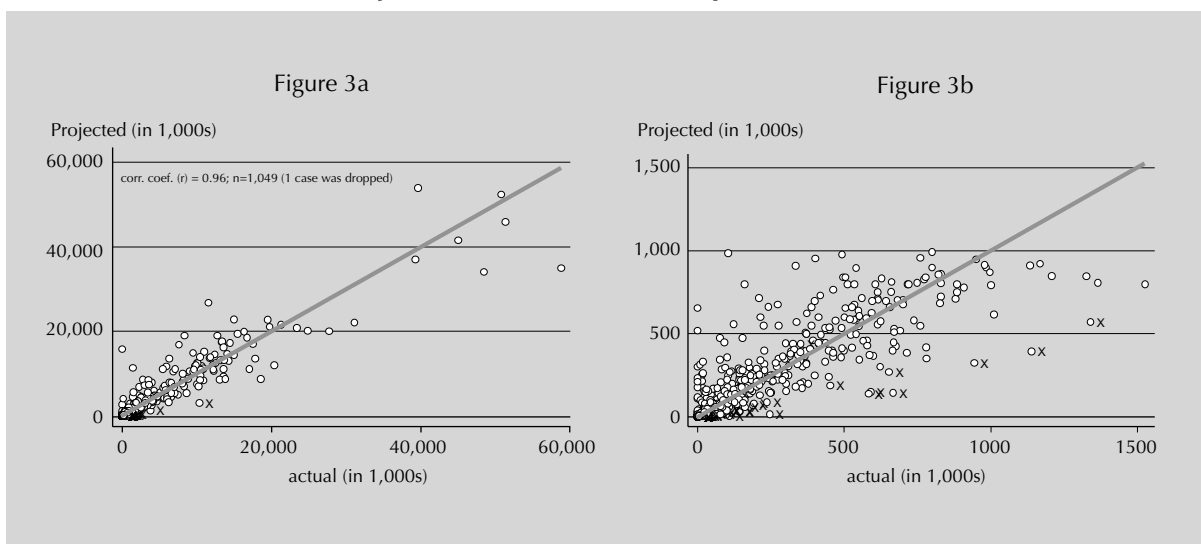
Figure 2.
Percentage Distribution of the Reported Quantity of Contraceptives Procured According to Donor, 1995–2003 Pooled



Assessing Outliers

Figure 3a displays a scattered plot between the projected and the actual contraceptive use pooled over the period 1995 to 2003. The dots that fall on the straight line in the center of the graph have no forecast error. However, forecasts often overestimate (as indicated by the dots above the line) or underestimate (as indicated by the dots below the line) the actual consumption as expected. The 47 outliers identified earlier were marked “X,” but they did not appear clearly in figure 3a because the outliers were associated with forecasts for smaller quantities of a contraceptive (located at the bottom left corner of the graph). To look closely at the

Figure 3.
Scatter Plots between the Projected and Actual Consumption, 1995–2003 Pooled



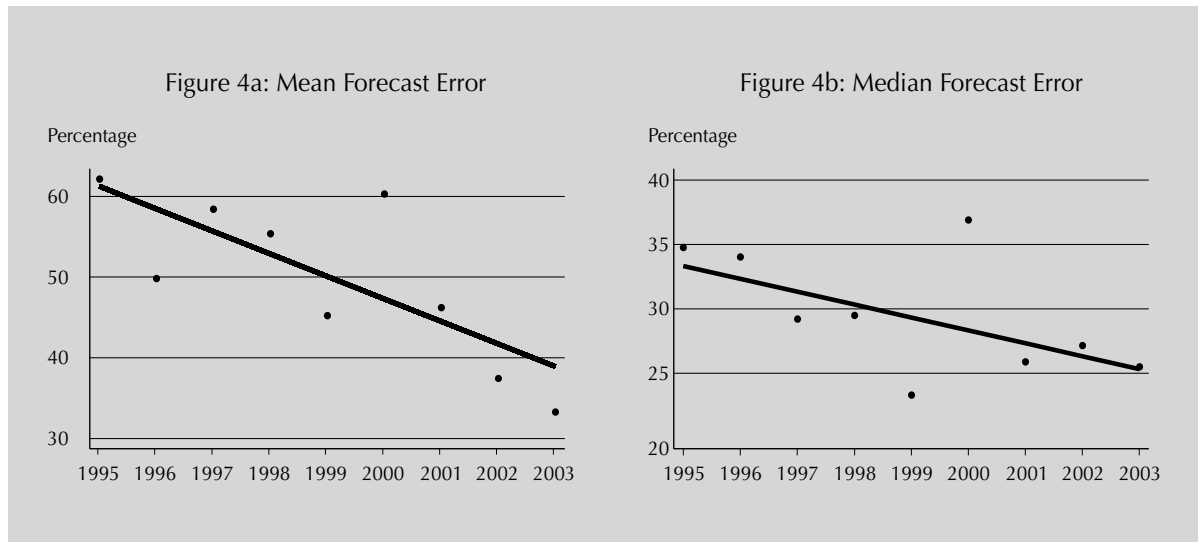
outliers, figure 3a was again plotted to create figure 3b but, this time, the sample was restricted to the contraceptive with a projected quantity of less than 1 million units. As expected, the outliers were clearer in figure 3b.⁵

Forecast Accuracy Over Time

Figures 4a and 4b, respectively, show the trend in the mean and the median forecast error or accuracy for all methods during 1995 to 2003. Each dot in the graph represents the mean or median forecast error for all methods for a given year. In 1995, the mean forecast error was 62 percent, and except for an outlier year in 2000 (when it was 60 percent), the mean forecast error steadily declined during the analysis period and reached its lowest error (33 percent) in 2003 (see figure 4a). Regression methods were used to test the statistical significance of the observed trend in the mean forecast error. After controlling for the variation of the mean forecast error over the analysis period due to country-level, product-level, and client-level differences, the regression model indicated that the observed declining trend was statistically significant⁶, and the mean forecast error declined by an average of about 4.4 percentage points per year.⁷ The straight line in figure 4a shows the declining trend of the mean forecast error.

Similarly, as expected, figure 4b shows that the trend of the median forecast error was consistent with the trend in the mean forecast error. Apart for the outlier year 2000 (for which the median forecast error was 34 percent), the median forecast error also steadily declined over time. In 1995, the median forecast error was 35 percent⁸ and by 2003 it had

Figure 4.
Trend in the Mean and Median Forecast Accuracy for All Methods, 1995–2003



5. The outliers appeared mainly as underestimates in figures 3a and 3b because of the way the forecast error was defined.

6. Alpha error or the p-value or the significance level for all statistical tests was set at 0.05. A p-value of less than 0.05 for the trend effect indicated that it was less than 5 percent chance that the observed declining trend in the mean forecast error was by chance. In brief, *statistically significant* is referred to as *significant* in the text.

7. All the average annual declines in the median forecast accuracy in this analysis were determined by the regression model using the marginal effect of trend; the results are included in the appropriate tables in the appendix.

declined to 26 percent. The regression model indicated that the observed trend in the median forecast error was statistically significant; on average, it declined by about 1.9 percentage points per year. The straight line in figure 4b shows the declining trend of the median forecast error. The methodological and numerical details of figures 4a and 4b are shown in table 2A and 1A, respectively.

It is noteworthy that the mean forecast error in all the forecast years is systematically higher than the corresponding median forecast error, indicating the presence of outliers and/or skewed distribution of the forecast error, which is consistent with what was discussed earlier. To avoid statistical complications⁹ from outliers and skewed distribution associated with the mean analysis, further discussion on the forecast error is based on their medians, which is robust to outliers and skewed distribution. However, the analyses of the mean forecast errors were also completed and reported in the appendix. By comparing the mean and the median analysis, it was possible to observe the consistency of the findings.

Forecast Accuracy by Method

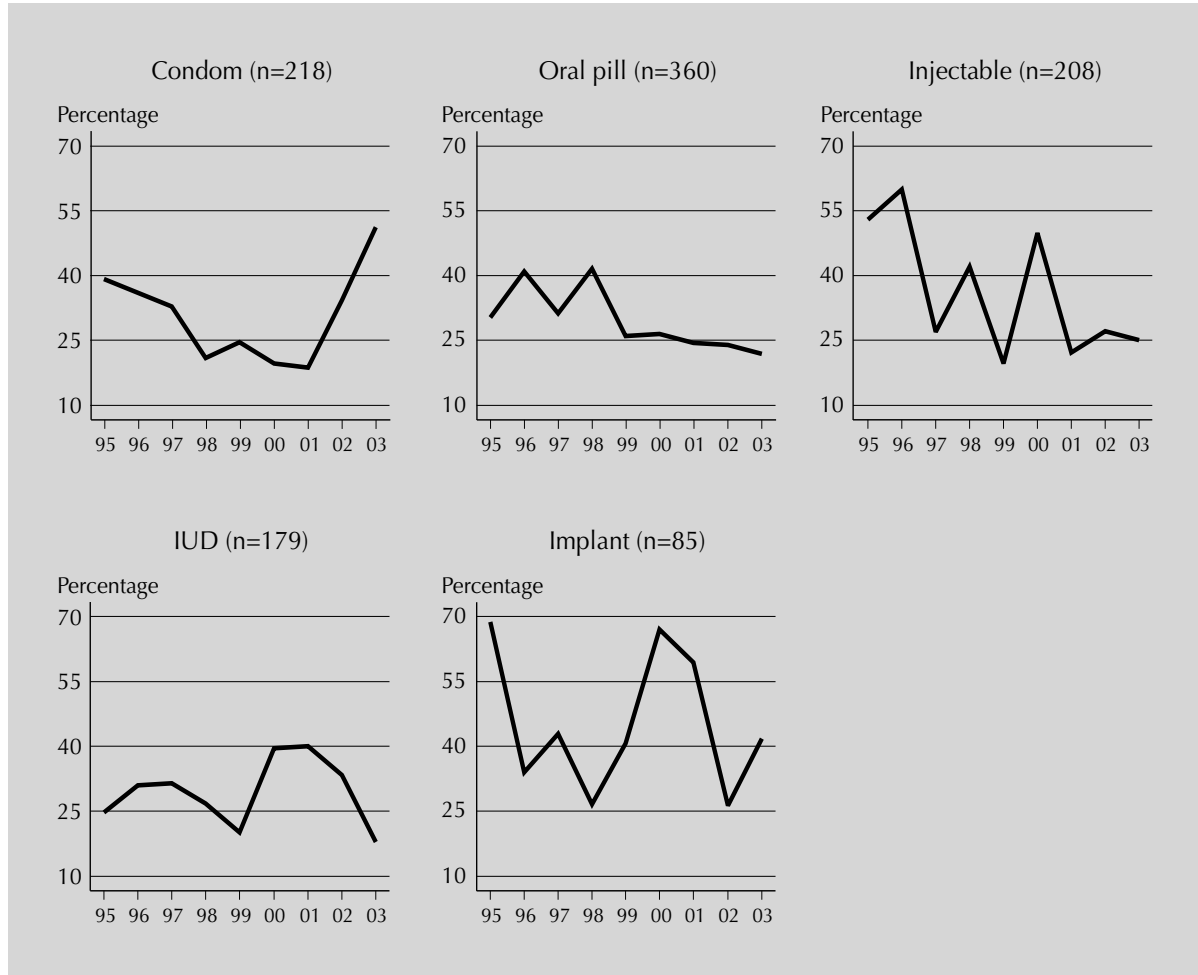
Figure 5 shows the trend in the median forecast error, by method, for 1995 to 2003. Regression methods were used to test the statistical significance of the trend effect of each of the methods. See table 1A for the numerical details for figure 5, including the description of the regression models that were used to test the trend effect. The median forecast accuracy for condoms significantly improved between 1995 and 2001—by an average of 3.7 percentage points annually. In 1995, the median forecast error for condoms was 39 percent, which declined to 19 percent in 2001. However, since 2001, the median forecast error for condoms has been significantly increasing on an average of 11.5 percentage points per year. In 2003, the median forecast error for condoms was 51 percent.

The trend in the median forecast accuracy for pills showed two peaks of decline during the years 1996 and 1998; nevertheless, the regression model indicated that the median forecast accuracy for pills was significantly improving—by an average of 2.6 percentage points per year. It improved from 30 percent in 1995 to 22 percent in 2003. The trend in the median forecast accuracy for injectable showed three peaks of decline during the years 1996, 1998, and 2000; nevertheless, the forecast accuracy for injectables was also significantly improving over the past nine years—by about 4 percentage points per year—as indicated by the regression model. In 1995, the median forecast error for injectables was 53 percent, which declined to about 25 percent in 2003. The two peaks in the forecast error for injectables observed during the years 1996 and 1998 corresponded with the two peaks in the forecast error observed shown for pills during the same years, indicating that it is likely the forecast estimates for both methods during the two years were partly influenced by the same factor.

8. Half of the forecasts during 1995 had a forecast error of 35 percent or less.

9. Skewed distribution of the mean forecast error would not bias the effect estimates of the regression model; however, the skewed distribution could produce biased standard errors leading to biased results of the statistical tests.

Figure 5.
Trend the Median Forecast Accuracy by Method, 1995–2003

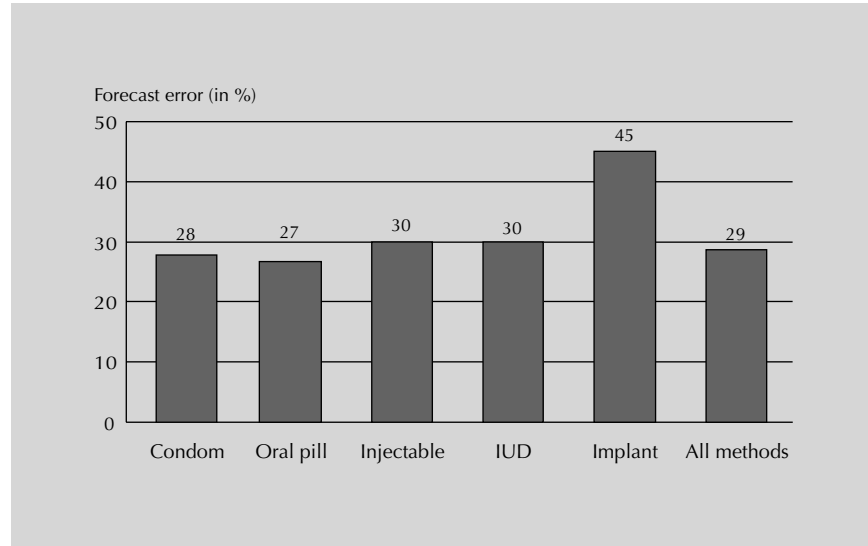


Except for 2000, 2001, and 2002, the median forecast accuracy for IUDs during the analysis period has been generally steady at about 30 percent or less. The trend in the forecast error for implants has been the very erratic. The highest median forecast error for implants were observed during 1995 (66 percent) and 2000 (67 percent); the lowest were observed during 1998 (27 percent) and 2002 (26 percent). The peak in the forecast error for implants, IUDs, and injectables observed during 2000 (see figure 5), in part, explains the peak in forecast error for all methods during the same year observed in figure 4b.

A statistical test was conducted to assess whether or not there were significant differences in the trend of the median forecast error between the different methods. As expected, the test indicated significant variation of the forecast accuracy trend by method.

Figure 6 shows the median forecast error by method, pooled over the period from 1995 to 2003. The highest median forecast error was observed for implants (45 percent), which were significantly higher when compared to the median forecast error for any of the other method.

Figure 6.
Median Forecast Error by Method, 1995 to 2003 Pooled



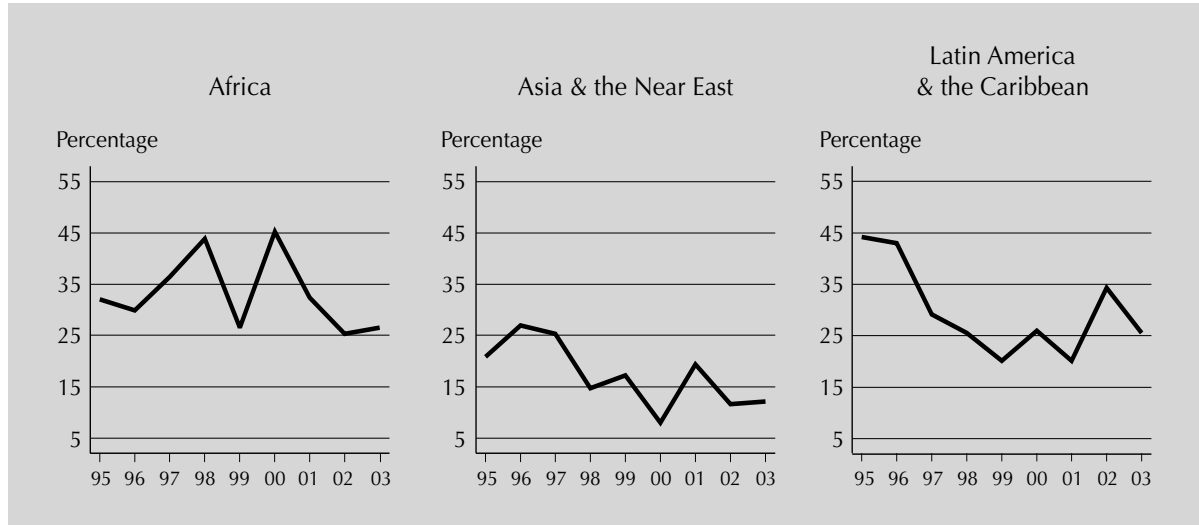
Regional Variation in Forecast Accuracy

Figure 7 shows the trend in the median forecast error for all methods, by region, from 1995 to 2003 (see table 1A for the numerical details). Throughout the analysis period, the median forecast errors observed in the ANE region were significantly less than those observed in the Africa region or the LAC region. The observed declining trend in the forecast error in the ANE region was not statistically significant, probably because the forecast error in the region was low to begin with; therefore, because the rate of decline was too small, the statistical test lacked the desired power¹⁰ to detect it. Although the trend in the median forecast error in Africa region showed two peaks of increase in 1998 (44 percent) and 2000 (45 percent), the regression model indicated that, on average, the median forecast error for the region has been significantly declining during the past nine years by about 1.7 percentage points per year; it declined from 32 percent in 1995 to 27 percent in 2003. Further analysis (not shown) revealed that the peak in forecast error in 1998 in the Africa region was mainly due to the peak in the forecast errors in the region for IUDs, injectables, and pills; while the peak in 2000 in the region was due mainly to the increase in the forecast errors for IUDs, implants, and injectables.

The decline in the median forecast error in the LAC region appeared to be faster compared with the decline in the other two regions. On average, the median forecast error in the LAC region significantly declined by an average of about 3.2 percentage points annually, from 44 percent in 1995 to 25 percent in 2003. The observed difference in the trend of the median forecast error between the three regions was not statistically significant.

10. The chance of the statistical test to detect the change in the forecast error given that the change actually occurred.

Figure 7.
The Trend in the Median Forecast Error by Region, 1995–2003



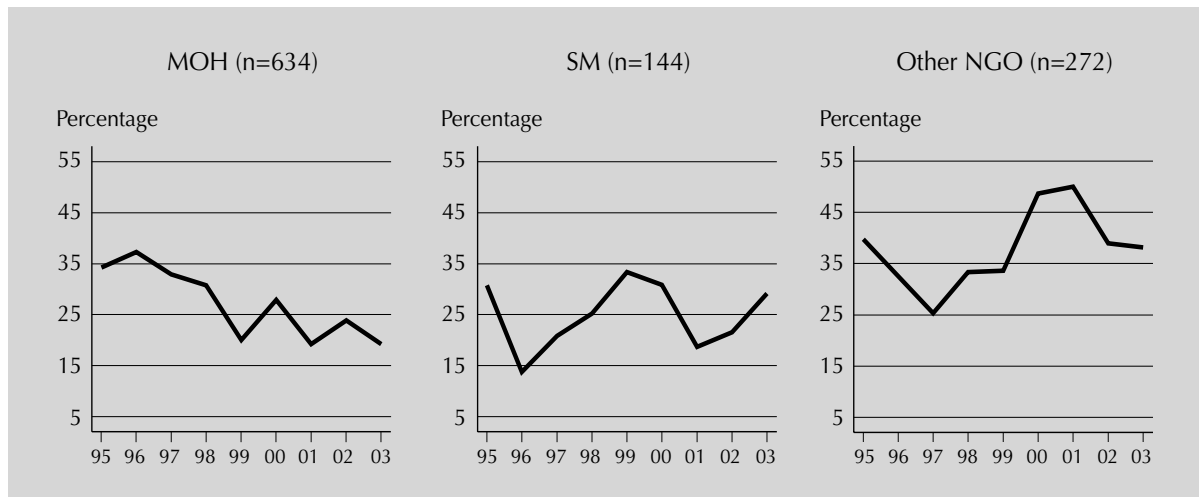
Variation of the Forecast Error by Client Category

Figure 8 shows the median forecast error, for all methods, by client category, pooled over the period 1995 to 2003. SM clients usually forecast contraceptive requirements based on past sales and future targets, which not necessarily account for the actual amount sold to users. Therefore, the forecast for SM clients are expected to be prone to larger error, while the projections for MOH clients are usually based on the historical or estimated dispensed-to-user data, which are expected to be less prone to forecast errors. Contrary to the expectation, figure 8 shows that the median forecast error for the SM clients (26 percent) was similar to those of the MOH clients (27 percent). However, the median forecast error for other NGO clients (40 percent) was much higher and was significant when compared to the SM or to the MOH clients.

Figure 8.
Median Forecast Error by Client Category, 1995 to 2003 Pooled



Figure 9.
Trend in the Median Forecast Error for all Method by Client Category, 1995–2003



Next, the trend in the median forecast error was analyzed for all methods, by client type, from 1995 to 2003 (see figure 9 and table 1A). The trend in the median forecast error for MOH clients significantly declined by about 2.1 percentage points per year, from 34 percent in 1995 to 19 percent in 2003. For SM clients, it was not conclusive if the observed trend in the forecast error was increasing or decreasing. The median forecast error for other NGO clients showed significant variation within the analysis period. From 1995 to 1997, the median forecast error for the other NGO clients declined from 40 percent to 25 percent. However, after 1997 it started increasing and, in 2001, reached a high of 50 percent. After 2001, the median forecast error for the other NGO clients again declined and reached 38 percent in 2003. As expected, the statistical test that assessed whether there was any difference in the observed trend of the median forecast error between the three client categories was significant.

Variation in the Forecast Accuracy by Number of Donors

Figure 10 shows the median forecast error polled over the period 1995 to 2003, for all methods, by the number of donors involved with supplying the contraceptives. It was expected that the excess administrative burden required to procure contraceptives using multiple donors upset the procurement plan, leading to a significant deviation between the actual and the projected use. However, the forecast accuracy was not significantly different between single (30 percent) and multiple donors (26 percent).

The trend in the median forecast error for all methods by number of donors is shown in figure 11 (see table 1A). The trend in the median forecast error showed a significant decline for single as well as multiple donors. The rate of decline in the median forecast error appeared to be faster for multiple donors (on average declining by 2.3 percentage points per year) compared to the rate observed for single donor (on average declining by 1.7 percentage points per year); however, the difference in the rate of reduction between single and multiple donor was not statistically significant.

Figure 10.
Median Forecast Error by Number of Donor, 1995 to 2003 Pooled

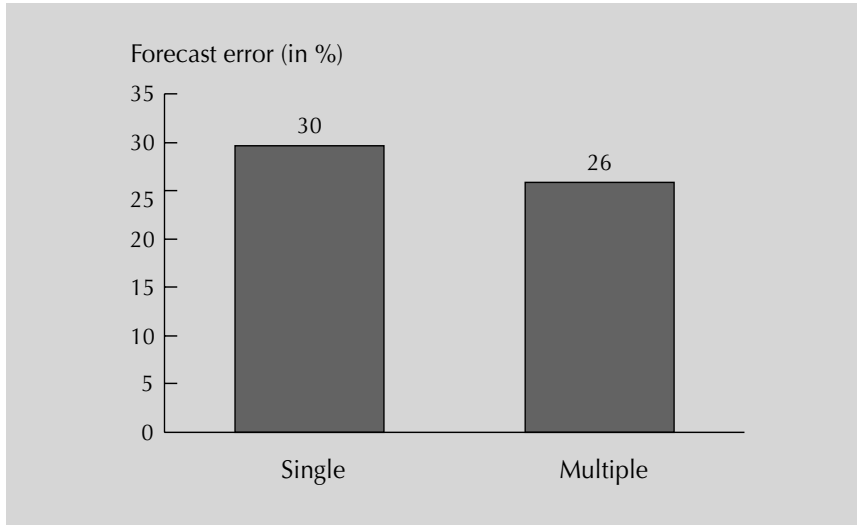
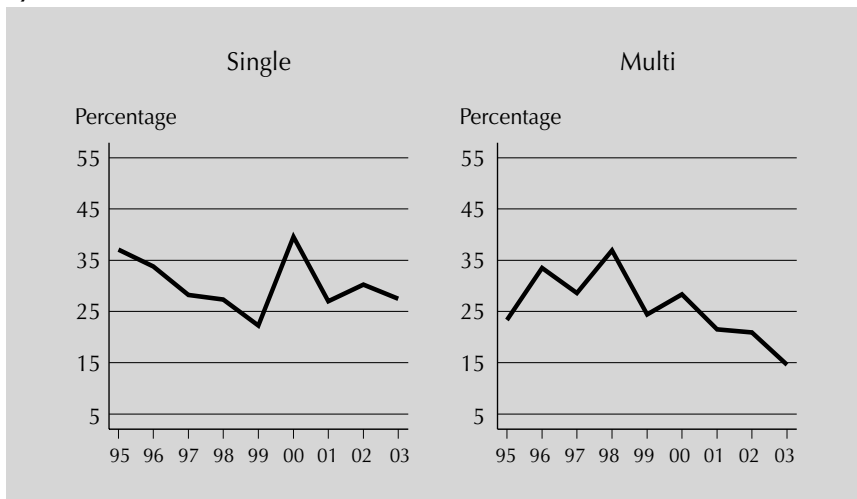


Figure 11.
Trend in the Median Forecast Error for All Methods by Number of Donors, 1995–2003



Variation of the Forecast Error by the Quantity Projected

The assessment of the outliers conducted earlier indicated that forecast error could be associated with forecasts for smaller quantities of a contraceptive. To confirm this, further analysis was done. To observe the relationship between forecast error and the quantity projected for a contraceptive, the projected quantity was grouped into five quintiles. The first quintile or the group that contained 20 percent of the sample with the smallest projected quantity of a contraceptive, had a range between 0.1 and 7 thousand; the next 20 percent of the sample or the second quintile ranged between 7.1 and 68 thousand. The middle quintile ranged between 68.1 and 350 thousand; the fourth quintile ranged between 350.1 and 1,548 thousand; and the largest quintile, or the group containing 20

percent of the sample with the largest projected quantity of a contraceptive, ranged from 1,548.1 to 166,000 thousand. The median forecast error by the five quintiles was estimated and then compared (see table 3). As expected, the analysis showed that the median forecast error was higher when the quantity projected for a contraceptive was smaller. For the smallest quintile, the median forecast error was 42 percent, which gradually decreased when the quantity projected increased; it reached 22 percent for the largest quintile. The observed relationship between forecast accuracy and the quantity projected was statistically significant and remained significant even after the variation of the forecast error due to client, country, product, and forecast year was controlled.

Table 3. Median Forecast Error by Quintiles of the Quantity Projected, 1995–2003 Pooled

Quintiles	Value range in each quintile (in 1,000s)	Median forecast error (in %)
Q1 (Smallest)	0.1 to 7	41.7
Q2	7.1 to 68	39.5
Q3 (Middle)	68.1 to 350	28.2
Q4	350.1 to 1,548	24.7
Q5 (Largest)	1,548.1 to 166,000	21.6

Mean Forecast Accuracy Analysis

The variation of the trend in the mean forecast accuracy by method, region, client category, and number of donors was also assessed. The findings of the mean forecast error analysis were mostly consistent with the median forecast analysis. See table 2A for the mean forecast accuracy analysis.

Forecast Accuracy by Country

The trend in the median forecast error for all methods was also analyzed by country (see table 3A). Significant declining trends in the forecast error were observed in Cameroon, El Salvador, and Tanzania. Although the median forecast error appeared as a declining trend in some of the other countries (e.g., Burkina Faso, Haiti, and Togo), the trend effect was not statistically significant because of scanty or missing observations for one or more of the forecast years. The median forecast error for Nepal, Philippines, and Zimbabwe did not show any specific trend during the analysis period; however, the forecast errors remained relatively low compared with the other countries. The average median forecast error during the analysis period for Nepal, Philippines, and Zimbabwe was 16, 11, and 19 percent, respectively.

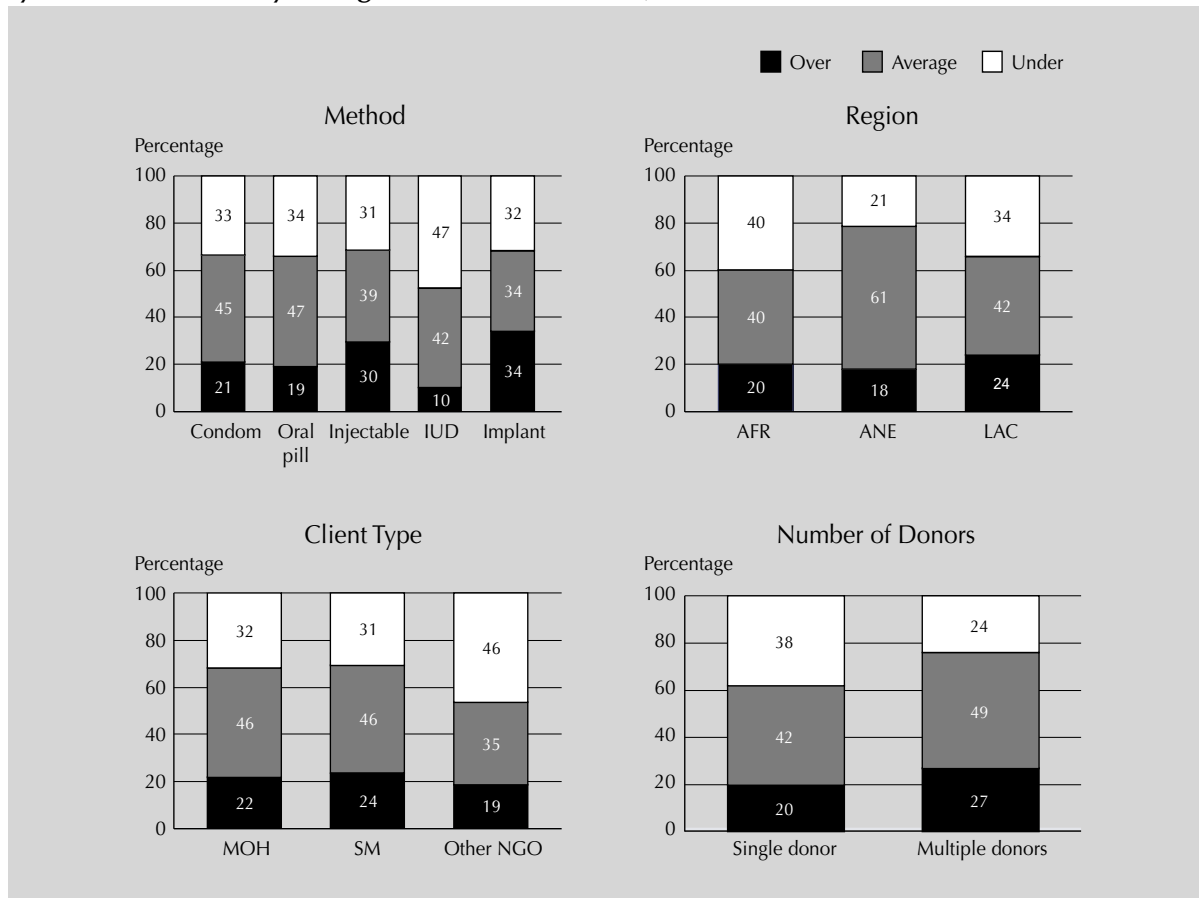
Direction of the Forecast Errors

The analysis completed so far provided information about the magnitude of the forecast error, but it was not clear whether the forecast error over-

or under-projected the actual consumption of a contraceptive. Therefore, to identify the extent to which the projected use overestimated or underestimated actual consumption, the percentage difference between the projected and actual consumption were categorized into three groups: (1) the projected use for a contraceptive that underestimated consumption by more than 25 percent; (2) the projected use that was within ± 25 percent of the actual use, referred to as the *average* in the table and figures for this section; and (3) the projected use that overestimated consumption by more than 25 percent.¹¹

The variation of the percentage of the pooled forecast that overestimated, was average, or underestimated actual consumption by method, region, client category, and number of donors was assessed (see figure 12 and table 4A). Figure 12 shows that forecasts are almost two times more likely to overestimate rather than to underestimate actual consumption. During the analysis period, on average, 35 percent of the forecast overestimated actual consumption, 21 percent underestimated it, and 43 percent were

Figure 12.
Percentage of the Forecast That Overestimated or Underestimated Actual Consumption by More Than 25% by Background Characteristics, 1995–2003 Pooled



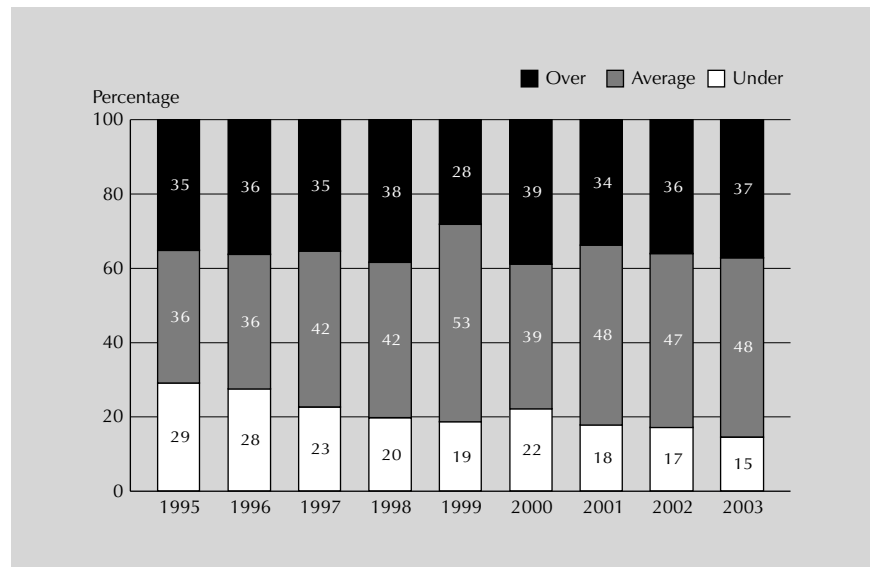
11. The cutoff point (i.e., ± 25 percent) for the average forecast error was decided based on the median forecast error for all method observed in 2003.

within ± 25 percent of the actual consumption. The over- or under-forecasting varied by method, region, client category, and number of donors. For different methods, the percentage of over-forecasting was higher for IUDs (47 percent) compared to other methods, which ranged between 31 and 34 percent. The percentage of under-forecasting was highest for injectables (30 percent) and implants (34 percent), and lowest for IUDs (10 percent). For different regions, the percentage of over-forecasting was lower in the ANE region (21 percent) compared to Africa (40 percent) or the LAC region (34 percent). For different client categories, the percentage of over-forecasting was higher for other NGOs (46 percent) compared to the MOH (32 percent) or SM (31 percent). And, for the number of donors, the percentage of over-forecasting was higher for single donors (38 percent) compared to multiple donors (24 percent).

Figure 13 shows the trend in the percentage of the forecast that overestimated, were within average rates, or underestimated actual consumption from 1995 to 2003, for all methods. Table 4A provides the details for figure 13. The trend in the percentage of under-forecasting declined significantly from 29 percent in 1995 to 15 percent in 2003, while the trend in the percentage within average forecasting increased significantly from 36 percent in 1995 to 48 percent in 2003. The improving trend in the percentage average for forecasting was consistent with the earlier findings that showed an improving trend in the mean and the median forecast accuracy. There was no significant difference in the percentage of over-forecasting the consumption between the forecast years. The observed improving trend in the forecast error for all methods is mainly contributed by the declining trend of the percentage of under-forecasting.

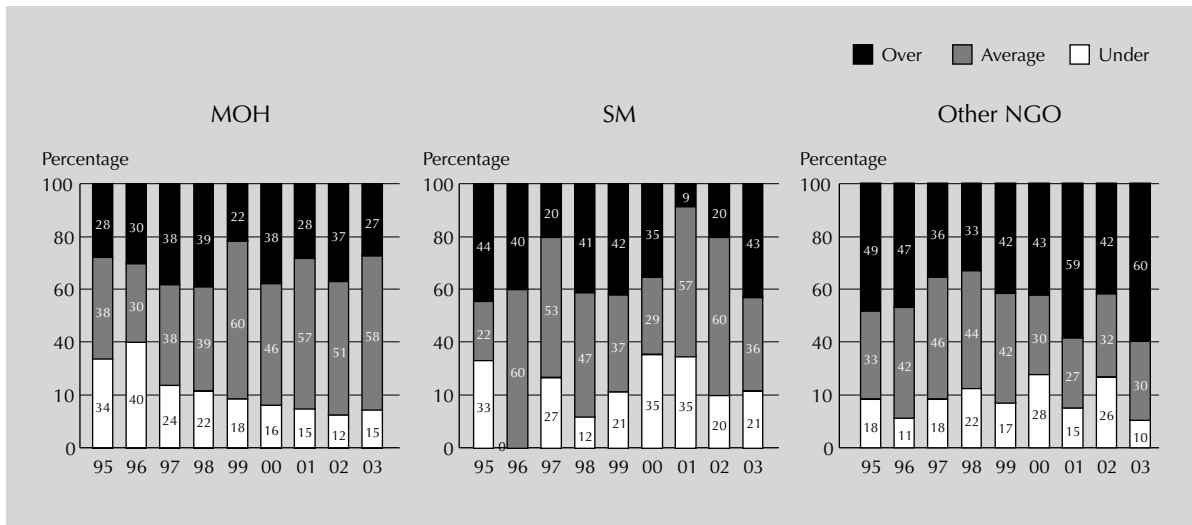
The authors assessed whether or not the trend of the direction of the forecast error varied by method, region, client, and number of donors (see table 4A). The analysis showed that the trend in the direction of the

Figure 13.
Trend in the Percentage Over- or Under-Forecasting, 1995–2003



forecast error varied significantly by client category and region (see figures 14 and 15, and table 4A). For the MOH clients, the trend (see figure 14) was similar to the overall trend (see figure 13), the percentage of under-forecasting for all methods declined significantly, from 34 percent in 1995 to 15 percent in 2003; while the percentage of the forecast within the average increased significantly from 38 percent in 1995 to 58 percent in 2003. The percentage of over-forecasting did not significantly vary between the forecast years.

Figure 14.
Trend in the Percentage Over- or Under-Forecasting by Client Category, 1995–2003

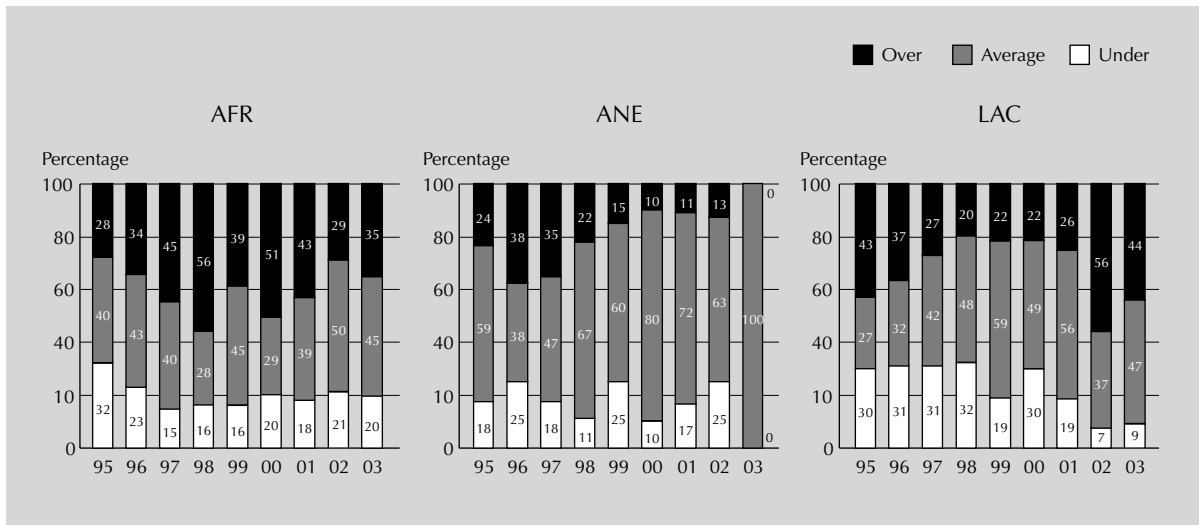


For the SM clients (see figure 14), the percentage of over-forecasting did not show any obvious indication of an increasing or decreasing trend. The percentage of over-forecasting for SM clients was 40 percent or higher during five of the nine forecast years (i.e., during 1995, 1996, 1998, 1999, and 2003). The expectation that the forecast for SM clients are more likely to overestimate consumption compared to the forecast for MOH clients was also true during the same five years. However, the percentage of over-forecasting for SM clients was lower during the forecast years 1997, 2001, and 2002 (20 percent, 9 percent, and 20 percent, respectively) compared to the lowest overestimated year observed for MOH clients (i.e., 22 percent in 1999). No significant trend was observed in the percentage of over- or under-forecasting among the other NGO clients, although over- or under-forecasting varied significantly between the forecast years.

Figure 15 shows the variation in the trend of the percentage of the forecast that overestimated or underestimated consumption for all methods, by region. In the Africa region, the percentage of over-forecasting did not show any obvious indication of an increasing or decreasing trend. Although, earlier the declining trend in the median forecast error for the ANE region was not found to be significant, the percentage of the forecast in the region that was within average showed an significant increasing trend, from an average of about 48 percent between 1995 and 1997 to an

average of about 73 percent between 2001 and 2003. This indicates an improvement in the trend of the forecast error for that region. The improving trend in the forecast error in the ANE region is mainly contributed by the declining trend of the percentage of the forecast of overestimating consumption, which is contrary to the overall trend observed in figure 13. The percentage of over-forecasting in the region significantly declined, from an average of 32 percent between 1995 and 1997 to an average of about 10 percent between 2000 and 2002; while the percentage of under-forecasting did not show any significant trend.

Figure 15.
Trend in the Percentage Over- or Under-Forecasting by Region, 1995–2003



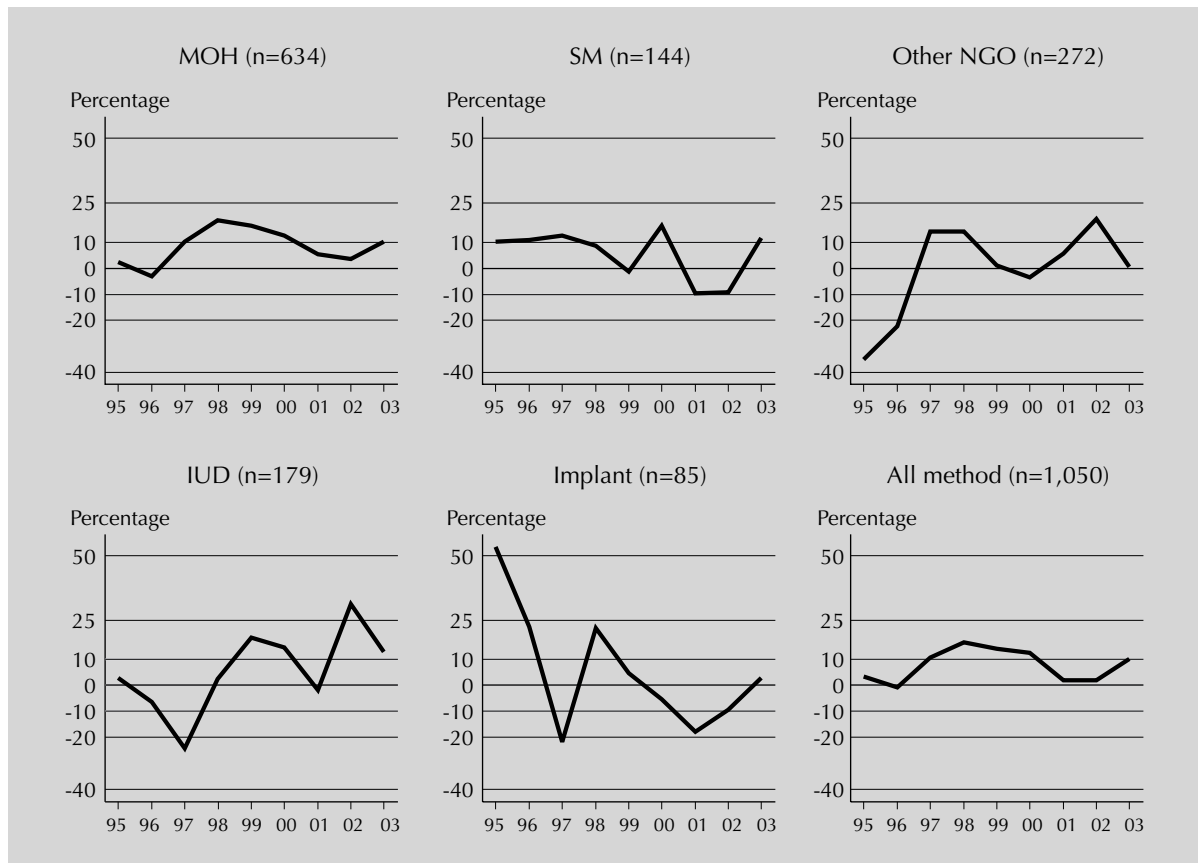
For the LAC region, the percentage of the forecast that significantly underestimated consumption declined from 30 percent in 1995 to 9 percent in 2003. Although the average forecasting in the region showed a decline during 2002, the overall trend in the average forecasting showed a significant increase, from 27 percent in 1995 to 47 percent in 2003. The percentage of over-forecasting in the region showed a significantly decreasing trend between 1995 and 2001; however, it increased again during recent years (i.e., 2002 and 2003).

It is interesting to note in table 4A that the recent increase (i.e., between 2001 and 2003) in forecast error observed for condoms earlier is mainly contributed by the increase in over-forecasting for the method. As expected, the significant trend effects of the percentage of the forecast within the average, by the different categories of background characteristics (see table 4A), was mostly consistent with the significant trend effect of the mean (in table 2A) and the median forecast error (in table 1A) for the same categories.

Aggregate Percentage Difference between Projected Use and Actual Use

To assess the utility of the CPTs for global procurement planning by CSL/USAID, all projected use and actual consumption were aggregated by background characteristics (i.e., method, region, client, and number of donors) to give the aggregated or the global projected and actual use. The percentage difference between the global projected and actual use was also estimated and reported in table 5A. Figure 16 shows the trend in the percentage difference between the aggregated projected and actual consumption of contraceptives, by method, from 1995 to 2003. For most of the projection years, the percentage difference between the global projected and actual use was positive, indicating the over-forecasting tendency at the aggregate level. For all methods, the trend in the aggregate percentage difference remained more or less stable at about positive 10 percent during the analysis period. No remarkable variation of the aggregate percentage difference for condoms and oral pills was observed during the analysis period. For injectables and implants, the aggregate percentage difference was remarkably high during the year 1995 (negative 35 percent for injectable and positive 53 percent for implant) compared to the other

Figure 16.
Trend in the Percentage Difference between the Aggregate Projected and the Aggregate Actual Consumption by Method, 1995–2003



years; it varied within about ± 20 percent throughout the rest of the analysis period, with the lowest aggregate percentage difference being observed during 2003 (1 and 3 percent for injectables and implants, respectively). The aggregate percentage difference for IUDs was comparatively high during 1997 (negative 24 percent) and 2001 (positive 31 percent) and comparatively low during 1995 (positive 3 percent), 1998 (positive 3 percent), and 2001 (negative 2 percent).

Figure 17 shows the trend in the aggregate percentage difference between the projected and actual consumption for all methods, by region, from 1995 to 2003. The trend in the aggregate percentage difference in the LAC region shows a linear shift from aggregate under-projection in the early part of the analysis period (between 1995 and 1996) to aggregate over-projection during the later part of the analysis period (between 2000 and 2003) with an indication of increase in the aggregate over-projection during recent years. The trend in the aggregate percentage difference in the Africa region was steadier compared to the trend observed in the ANE or the LAC region. No definitive trend in the aggregate percentage difference was observed for the ANE region.

Figure 17.
Trend in the Percentage Difference between the Aggregate Projected and the Aggregate Actual Consumption by Region, 1995–2003

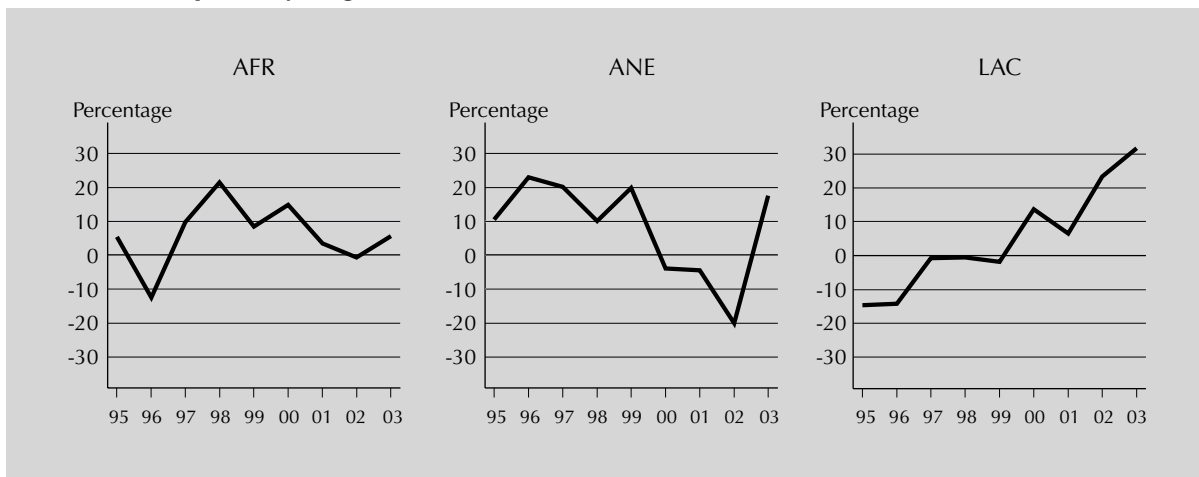
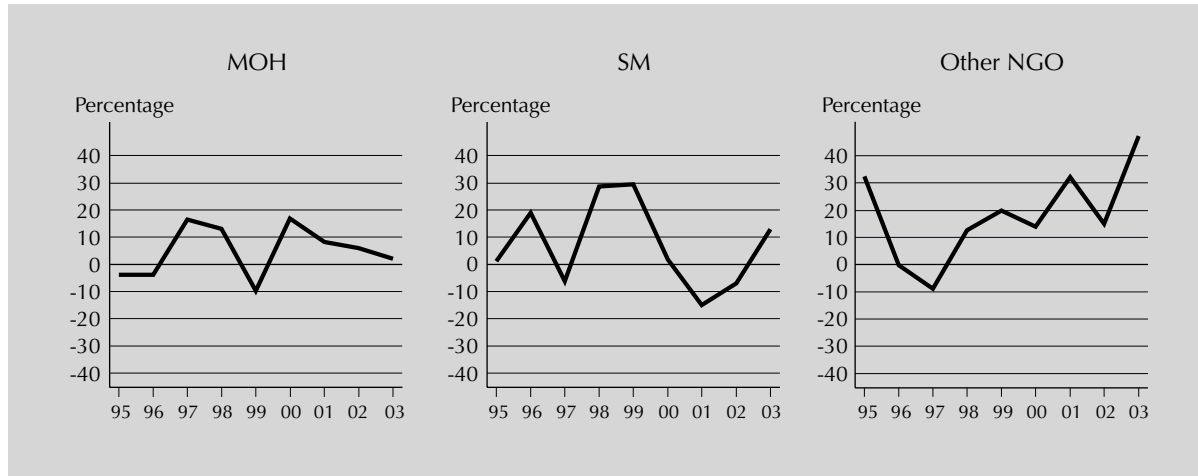


Figure 18 shows the trend in the aggregate percentage difference between the projected and actual consumption for all methods, by client category, from 1995 to 2003. The variation in the aggregate percentage difference within the analysis period was lowest for MOH clients compared to the SM or to other NGO clients. The trend in the aggregate percentage difference between 1998 and 2003 for other NGO clients shows an gradual increase, from positive 13 percent in 1998 to positive 47 percent in 2003. No definitive trend in the aggregate percentage difference was observed for the SM clients.

Figure 18.
Trend in the Percentage Difference between the Aggregate Projected and the Aggregate Actual Consumption by Client Category, 1995–2003



Comparison between Projected Shipment Accuracy and Forecast Accuracy

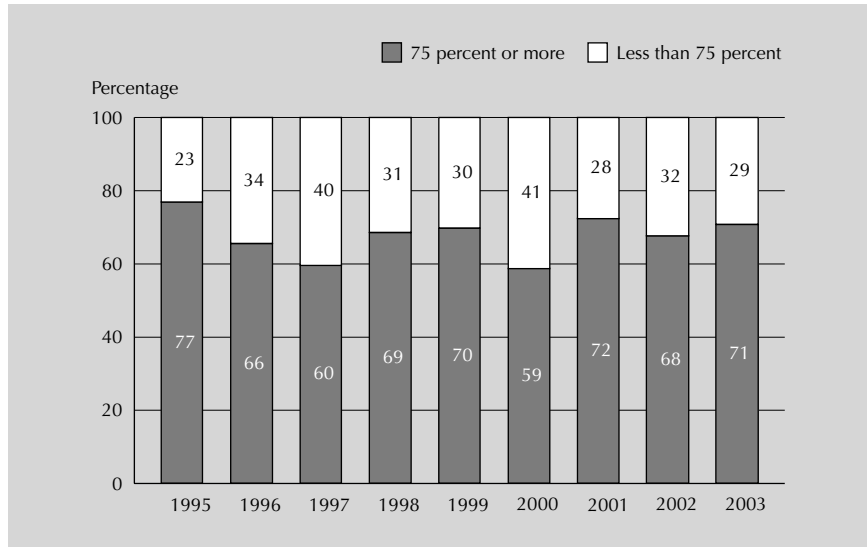
The authors assessed the relationship of the variance between the proposed and actual shipment, i.e., the projected shipment accuracy, and the variance in projected and actual use of a contraceptive. The projected shipment accuracy was defined in terms of its adequacy. The proposed quantity of a particular brand of contraceptive for the CPT planning year is compared to the actual quantity of the product received by the client and obtained from the two-years-later CPT. If a client or program received 75 percent or more of the quantity proposed for shipment for a particular product, then the client or program is considered to have adequate projected shipment accuracy.¹²

Figure 19 shows the trend in the percentage of the cases that received 75 percent or more of what was planned for shipment for a contraceptive, i.e., the trend in receiving an adequate shipment, from 1995 to 2003. The percentage having an adequate shipment did not vary by much during the analysis period. The highest percentage (77 percent) having adequate shipment was observed during 1995, while the lowest was observed during the 1997 (60 percent) and 2000 (59 percent). The authors assessed the variation of the percentage having adequate shipment by the background characteristics (not shown); however, it did not reveal any significant findings.

The authors assessed the relationship between the direction of the forecast error and receiving adequate shipment. Figure 20 shows the distribution of the direction of the forecast error by adequate or not adequate shipment status, pooled over the period 1995 to 2003. The analysis shows a statistically significant relationship between the direction of the forecast

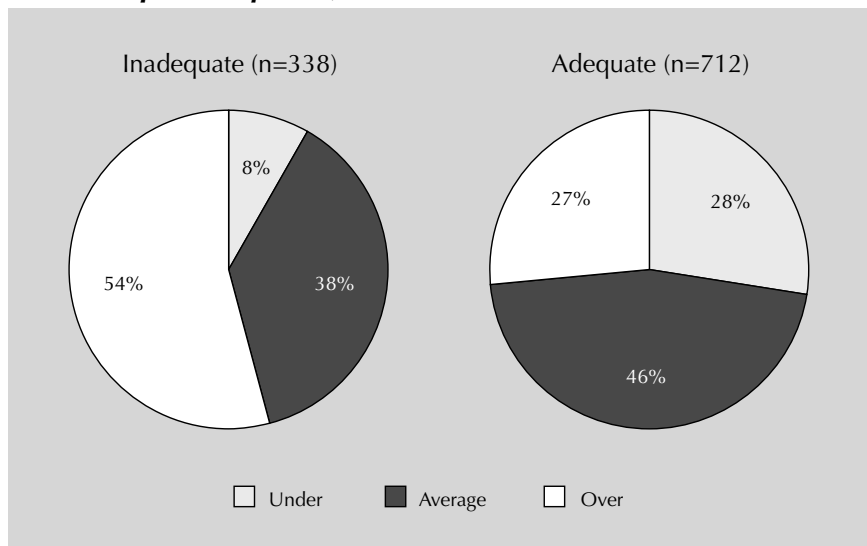
12. Percentage difference between projected and actual shipment = $100 \times (\text{projected shipment} - \text{actual shipment}) \div (\text{projected shipment} + 0.01)$. The constant '0.01' was added to the denominator of the shipment accuracy measure so the CPTs that showed zero quantity planned for the forecast year could be defined, and therefore, not excluded.

Figure 19.
Trend in the Percentage of the Proposed Shipment Realized in Actual Shipment by 75% or More in Quantity, 1995–2003



accuracy error and the shipment receiving status in the expected direction. Overestimation of the forecast error was twice as high when there was not an adequate shipment (54 percent) compared to when there was adequate shipment (27 percent) during the forecast year. The percentage of the forecast error within the average was higher when the shipment status was adequate (46 percent) compared to when the shipment status was not adequate (38 percent).

Figure 20.
Percentage Under-, Average-, or Over-forecasting by Adequate and Inadequate Shipment, 1995–2003 Pooled

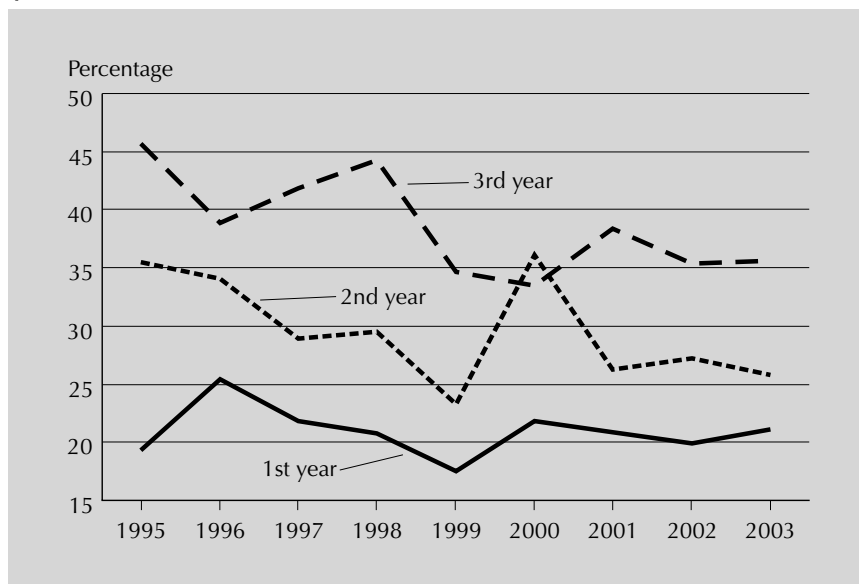


The significant relationship between the direction of the forecast error and shipment status remained even after accounting for some of the other sources of variation (i.e., controlling for country, client, and product, the analysis is not shown). One explanation of the observed relationship could be that an inadequate shipment causes lower consumption and leads to over-forecasting. Another explanation for the observed relationship could be that the CPT advisors know their client would order less than what was proposed; therefore, to ensure enough supply, they deliberately proposed a higher quantity to ship.

Comparison of the Forecast Accuracy within a CPT

This analysis was confined to validate the forecasts done for the CPT planning year. However, as discussed earlier, CPTs also contain a forecast for the CPT year and for the year following the CPT planning year. Next, a variation of the forecast accuracy was compared for the CPT year, for the CPT planning year, and for the year following the CPT planning year next. Forecast accuracy for the projected use for the CPT year and for the projected use for the year following the CPT planning year was estimated using the actual use obtained from one-year-later and three-year-later CPTs, respectively. Figure 21 compares the trend in the median forecast error by the three forecast years of a CPT. As expected, the longer the forecasts into the future, the higher the forecast error. The forecast accuracy for the CPT year was the most accurate, followed by the forecast accuracy for the CPT planning year, and the worst forecast error was observed for the year following the CPT planning year.

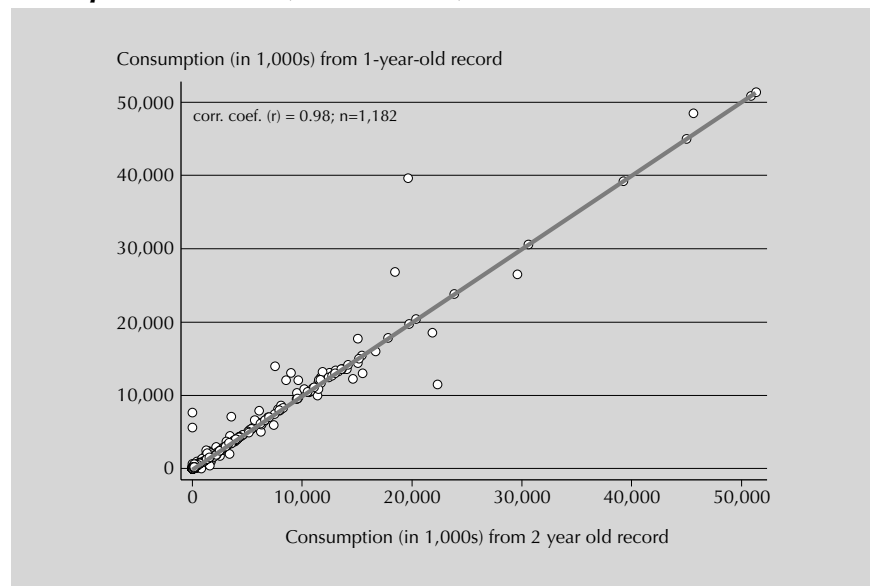
Figure 21.
Trend in the Median Forecast Error by 'CPT year' (year 1), 'CPT Planning Year' (year 2), and the Year after 'CPT Planning Year' (year 3), 1995–2003



Reliability of the Actual Consumption Reported in the CPTs

Although the validity of the actual consumption could not be assessed for this study, its reliability was assessed. The actual consumption during the year before the CPT year is again reported in the following year as the actual consumption two years before the CPT year, giving an opportunity to assess the reliability of the actual consumption reported. Because the projected use of a contraceptive is often estimated based on past consumption, it is expected that logistics advisors will review past consumption data reported in the earlier CPT and update it in the later CPT so that better consumption data is available to improve the forecasting. Figure 22 shows a scattered plot between the actual consumption reported between two subsequent CPT years for all method pooled over the analysis period. The straight line on the graph indicated 100 percent reliability for the actual consumption. The analysis indicated that the information on the actual consumption, in most cases, was reliable. More than half (56 percent) the dots fell on the straight line, while 77 percent of the cases showed reliability of 90 percent or higher.

Figure 22.
Scatter Plot between the Actual Consumption Reported within Two Subsequent CPT Years, All Methods, 1995–2003 Pooled



Discussion

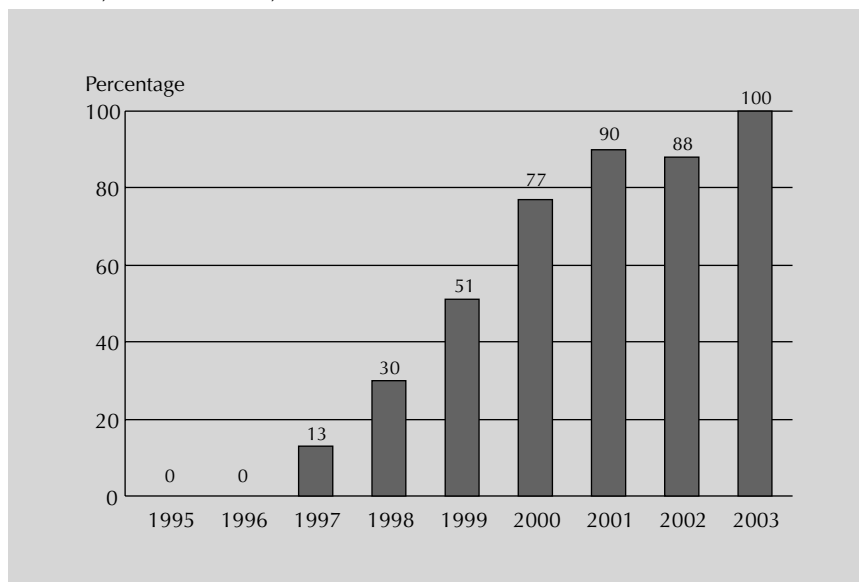
This analysis examined a selected number of CPTs between 1994 and 2004 contained in the NEWVERN database, which is maintained by DELIVER, to observe the trend and determinants of the forecast accuracy from 1995 to 2003. The selected sample represented 50 clients from 19 countries where CPTs were prepared using DELIVER’s input. After minimizing bias, the analysis found forecast accuracy improved over time. In general, the tendency of the projected use was to overestimate the

actual consumption. The pooled analysis of the forecast accuracy measures indicated that forecast accuracy varied by method, region, client category, and amount of contraceptive projected. The trend of the forecast accuracy varied by contraceptive method, client category, and region. The findings were consistent across the three different analytic methods that were used to assess the forecast accuracy.

The improvement of the forecast accuracy could be explained by one or more of the following: (a) improvement in the forecasting methodologies and procedures followed by the logistics advisors, (b) improvement in the ability of the CPT clients to obtain historical dispense-to-user data through an improved logistics management information system (LMIS), and (c) increase in the frequency of forecasting. The analysis period for this study reflects the project life of the Family Planning Logistics Management III (FPLM III) project and the first two-and-a-half years of the DELIVER project. One noteworthy improvement in the forecasting methodology during this period was the introduction and use of PipeLine software. In 1997, 13 percent of the forecast included in this analysis was done using PipeLine; the use gradually increased to 100 percent by 2003 (see figure 23). PipeLine facilitated the efficient preparation of CPTs through automation. It could have improved forecast accuracy by providing the means to estimate the actual consumption of a contraceptive with consistency and by avoiding mathematical errors.

The software also allowed estimating more frequently the projected quantity of a contraceptive needed by the client. However, without an effective LMIS in each country/program to feed the PipeLine with the appropriate information, it is unlikely that the implementation of the software alone could explain all the observed decline in the forecast error.

Figure 23.
Trend in the Percentage of the Cases That Used PipeLine to Prepare the CPT, All Methods, 1995–2003



An earlier analysis by Gelfeld (2000) showed that between 1995 and 2000 there was an improvement in the LMIS of the countries where FPLM III provided technical assistance, indicating that the declining forecast error during that period could be partly explained by the improvement in the logistics system. However, further investigation will be required to assess the situation.

The recent forecast accuracy observed in this analysis sets a standard accuracy level for all CPTs in the future. The level of the median forecast error during 2002 and 2003 was 26.7 percent with the 95 percent confidence interval ranging between 23.1 and 30.5 percent (see table 4). This is within the acceptable range (i.e., within 25 to 30 percent, on average) for a one-year-ahead forecast by the U.S. commercial forecasting standards (Wilson 1995).

Table 4. Forecast Accuracy at 25th, 50th, and 75th Percentile, All Methods, 2002 and 2003 (n=228)

	Forecast Accuracy (%)	95% Confidence Interval	
		<i>Lower limit</i>	<i>Upper limit</i>
25th percentile	13.9	10.9	16.9
50th percentile	26.7	23.1	30.6
75 th percentile	44.1	41.3	54.6

It is a common practice for countries with multiple family planning programs to conduct forecasting and procurement plan separately, even for a contraceptive that is supplied by the same donor. Since lower accuracy is associated with forecasting smaller quantity of a contraceptive, splitting up the forecasting of the national requirement for a contraceptive (into smaller quantities) by different programs may introduce error. This may be avoided by conducting pooled forecasting for contraceptives for all programs in a country.

Although the aggregate forecast for a contraceptive in the CPTs slightly overestimates the aggregate consumption, keeping the bias in mind, the aggregate forecast can be used by CSL/USAID to plan and monitor central level procurement planning.

The causal relationship between shipment accuracy and forecast accuracy was not conclusive, because both indicators were measured at the same point in time. Further analysis of the PipeLine database will be required to find the causal relation between the two indicators.

Chapter 2

Influence of PipeLine and LMIS on Forecast Accuracy

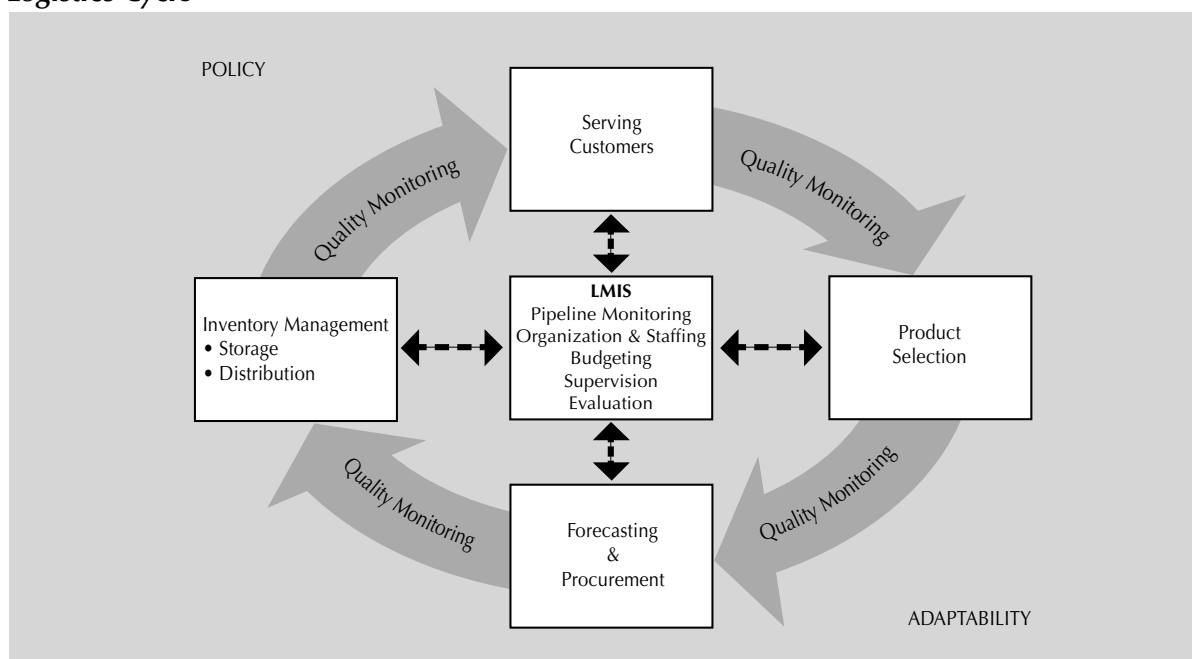
Introduction

The analysis in the previous chapter indicated that from 1995 through 2003 the forecast accuracy of the CPTs has improved. It has been suggested that the implementation of PipeLine software and the improvement of the logistics management capacity of the CPT clients have contributed to the improving trend in the forecast accuracy. The analysis in this chapter seeks empirical evidence to support the theories generated in the previous chapter by answering the following two research questions: (1) Has the implementation of PipeLine software resulted in the improvement of the forecast accuracy? and (2) Is the improving trend in the forecast accuracy due to the improvement of the CPT client's LMIS.

Analytic Framework

The analytic framework for this study is derived from the logistics cycle developed by the DELIVER project in 2003. Forecasting for a contraceptive is directly or indirectly influenced by all the components of the contraceptive logistics management system (see figure 24). However, the authors mainly

Figure 24.
Logistics Cycle



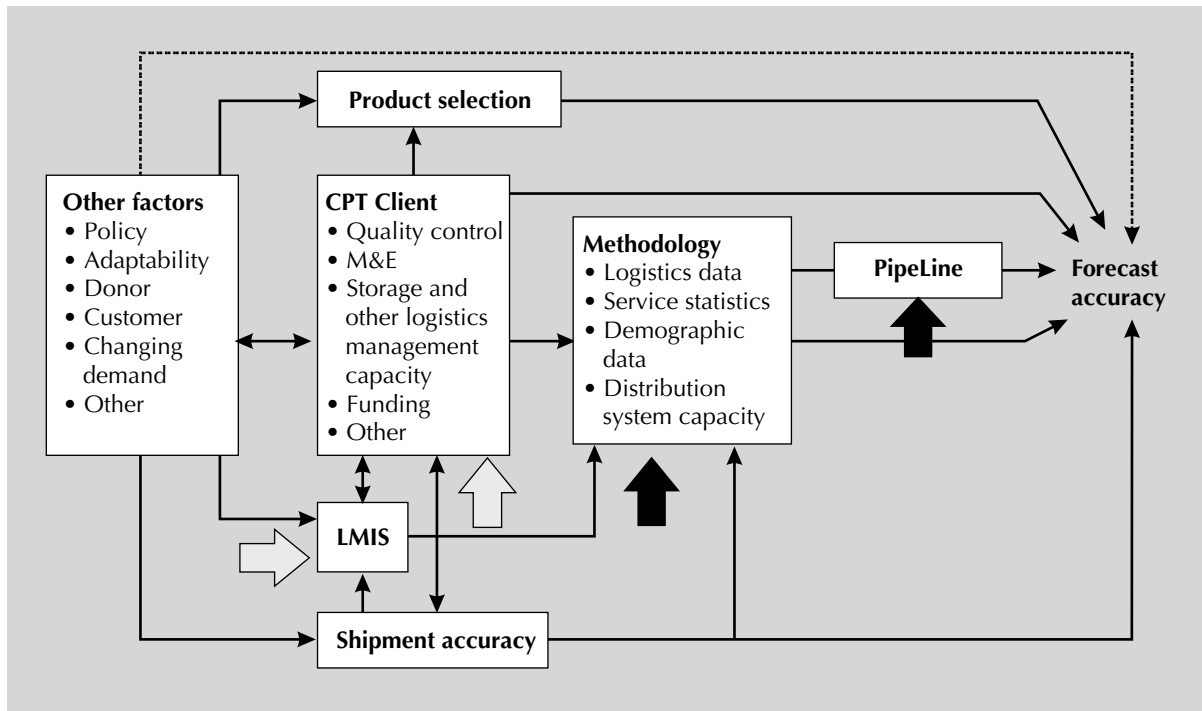
Source: JSI/DELIVER 2004.

focused on the components of the logistics system that are more proximate to the forecasting process for a contraceptive. The proximate factors that influence the accuracy of a one-year-ahead forecast for a contraceptive considered for this study include the selection of the product, CPT client characteristics, forecasting methodology, information system to support the methodology, and shipment accuracy.

Figure 25 shows the schematic diagram of the analytic framework for this study. Contraceptive forecasts using past consumption data from logistics data are considered to be the most accurate (FPLM 2000a). Contraceptive forecasts from logistics data are usually combined with one or more of the other forecasting methods (i.e., service statistics, demographic data, and distribution system capacity) to prepare a CPT. Therefore, it is expected that the more reliable the client's LMIS, the more accurate the forecast. The use of PipeLine software to conduct forecasts, on the other hand, is expected to improve the forecast accuracy by avoiding mathematical errors, maintaining consistency in estimating actual consumption, and allowing more frequent forecasting.

Product selection directly influences forecast accuracy because it is based on what the family planning user (the customer) prefers to choose and obtain; this, therefore, affects future consumption. The CPT client or the family planning program directly and indirectly influences the forecast accuracy by ensuring the maintenance of the logistics cycle. Delay and inadequate shipments of a contraceptive may lead to stockouts and, therefore, will reduce

Figure 25.
Analytic Framework



Note: Prominent arrowheads indicate FPLM/DELIVER's intervention; the transparent arrows indicate variable or partial intervention of the project.

the level of expected consumption. The CPT client plays a vital role in ensuring an adequate and timely supply of a contraceptive by ordering the required quantity of the contraceptive in a timely manner. However, external factors (the donors and manufacturers) may also contribute to unusual shipment delays or the ordering of inadequate quantities of a contraceptive. Therefore, the delay and inadequate shipment of products, or the shipment accuracy, may also have an independent influence on the forecast accuracy.

Other factors (e.g., policy environment, donor environment, family planning customers, exogenous factors changing demands, and other unmeasured factors) that may influence forecast accuracy are unaccounted in this study and are indicated by the dashed line in figure 25. However, most of the influence of the unaccounted or the unobserved factors are expected to exert their influence on forecast accuracy by acting through the more proximate determinants.

The interventions of DELIVER and its predecessor FPLM directly influenced the forecasting methodology and the implementation of PipeLine software; they appear as prominent arrowheads in figure 25. DELIVER and FPLM also provided technical assistance and training to CPT clients to improve the other aspects of the logistics cycle, including the LMIS. However, the degree of the intervention is expected to vary between clients and it remained unmeasured in this study. The unobserved or unmeasured influence of DELIVER and FPLM appear as transparent pointers in figure 25. The objective of this analysis is to demonstrate the impact of PipeLine and the LMIS on forecast accuracy, after controlling the influence of the other variables.

Data, Measurements, and Analytic Technique

Three different data sources were linked together for this analysis. These included the database used for the forecast accuracy analysis in the previous chapter, the PipeLine database maintained by the Central Contraceptive Procurement (CCP), DELIVER; and the composite indicator score sheet (CISS) database maintained by FPLM III and archived by DELIVER.

Dependent Variable

The forecast accuracy is the outcome or the dependent variable of this study. As described earlier, forecast accuracy is defined as the absolute percentage difference between the projected and actual use of a contraceptive. Three types of indicators are used to describe the forecast accuracy, including the median forecast error, the mean forecast error, and the percentage of the forecast within ± 25 percent of the actual consumption (also referred to as *within average forecasting*). The issues related to the measurement error of the mean and median forecasting was described in the previous chapter. The percentage within average forecasting is included as another indicator to describe the forecast accuracy because it measures whether or not a forecast for a contraceptive is within a set standard (i.e., within 25 percent of the forecast accuracy). The indicator being a binary response variable, it does not have the complication of a skewed distribution and robust to extreme

forecast inaccuracy (i.e., outliers). However, because the indicator does not indicate the degree of forecast inaccuracy, it is not very useful as a stand-alone measure of the forecast accuracy.

Independent Variables

Two independent variables were of major interest for this study, including the variable measuring of the implementation of the PipeLine software and the variable measuring of the level of the functional aspect of the LMIS of a program. Four other independent variables were included in this study as controls; these were product brand (other than for condom), country, client, and shipment accuracy. The definition for shipment accuracy was defined similarly to the definition in chapter 1. If a client or program received 75 percent or more of the quantity proposed for shipment for a particular product, then the client or program was considered to have adequate shipment accuracy.

PipeLine Implementation

The PipeLine database contained the name of the client and the date when the logistics data was first entered, indicating the date on which a particular client started using the software. The PipeLine database was linked with the forecast accuracy analysis database and a dichotomous response variable was created, which indicated whether a client in a given year implemented the PipeLine software to prepare its CPTs. The use of PipeLine in preparing a CPT was expected to be associated with a lower level of forecast error.

Functionality of LMIS

The extent to which the LMIS of a CPT client is operating was measured using the tool for Composite Indicators (CI) for Contraceptive Logistics Management, designed by the EVALUATION project and FPLM (FPLM 1999). The tool uses a structured questionnaire with 23 items to obtain information about eight different aspects¹³ of the logistics system of a family planning program by interviewing key informants. Each item is scored twice using a three- or five-point Likert-type scale response: once for measuring the performance and once for measuring the sustainability of a given aspect of the logistics system. The performance section of an item addressed, “How well is the logistics functioning?” The sustainability addressed, “How independent from donor support is the system?” The items were weighted and then aggregated separately to give the CI for performance and the CI for sustainability of a logistics system.

For its evaluation, FPLM III implemented the CI tool on 64 family planning programs in 28 countries where it provided technical assistance (Gelfeld 2000). The tool was once implemented in 1995, once in 1999, and once in 2000; the individual scores of the items of the CI indicator were main-

13. The eight aspects or components of the logistics system included LMIS, forecasting, procurement, warehousing, distribution, organization and staffing, policy, and adaptability.

tained in the CISS database. For this study, the CISS database was linked with the database created for the forecast accuracy analysis for the forecast years 1995, 1999, and 2000 by using the client- and country-level information. Altogether, the individual item scores for the CI from 26 programs in 14 countries could be linked with 207 cases (58 percent) of forecast accuracy measurements for the reference period. The scores of the four items on LMIS performance were aggregated to create an index that measured the functionality of the CPT client's LMIS. The minimum and maximum possible score of the index ranged between 0 to 12, and a comparatively higher score indicated the client had a comparatively better LMIS. A dose-response relationship was expected between the functionality of the LMIS and the forecast error, i.e., comparatively high score of the LMIS index was expected to be associated with a comparatively high forecast accuracy. See table 5 for a description of the items used to construct the LMIS index.

The CI is criticized for its subjectivity of the respondents, inter-rater reliability, and variance of the quality and source of data (Gelfeld 2000). Part of the subjectivity and reliability of the CI was improved by focusing only on the items related to the LMIS performance that are more tangible, compared to the other aspects (e.g., policy, adaptability, etc.) of the logistics system.

Analytic Technique

The effect of PipeLine and the LMIS on the forecast accuracy was analyzed separately due to the smaller sample available to analyze the effect of the latter. The effect of PipeLine was assessed using all 1,050 cases of the forecast accuracy measures between 1995 and 2003. The effect of LMIS on forecast accuracy was assessed using 207 cases of the forecast accuracy measures for 1995, 1999, and 2000.

The three different indicators of forecast accuracy were analyzed using three different statistical models, as appropriate. The median forecast error was analyzed using median regression methods with robust stan-

Table 5. Description of the Items Used to Construct the LMIS Capacity Index

Item	Max. Score	Mean Score		
		1995	1999	2001
Performance				
Program has basic elements of LMIS	4	1.8	3.2	2.9
LMIS information is used in management decision making	4	1.5	3.0	2.8
LMIS information is fed back to all lower levels in the distribution system	2	0.6	1.3	0.9
Commodities data are validated by cross-checking with other data sources	2	0.5	1.5	1.1
TOTAL Score for the LMIS capacity index	12	4.7	8.8	7.7

standard errors¹⁴; the mean forecast accuracy was analyzed using fixed-effects ordinary least square (OLS) methods; while the likelihood that the forecast was within the average was analyzed using fixed-effects logit methods. All three types of analysis were implemented using Stata (StataCorp 2003). The fixed-effects analogue of the median regression model was achieved using dummy variables for country, client, and product brand in the model as control variables (Wooldridge 2003; Koenker 2004).

The fixed-effect models accounted for the repeated measures by holding the effects of client, country, and product on the outcome constant, over time. The fixed-effects models also accounted for the portion of the measurement error on the dependent (i.e., forecast accuracy), as well on the independent variables (i.e., specifically the LMIS index and shipment accuracy) that remained constant over time. For example, the portion of the measurement error of the forecast accuracy due to the error in estimating the actual use would be eliminated by the method if they were similar between two points in time; the common error between time would difference out. However, the time varying measurement errors of the independent variables that are related with the outcome, remained as a threat to the validity of this study. A time varying measurement error of the forecast accuracy could happen if the subjective nature of the raters changed over time. For example, if the raters knew that the forecast accuracy was improving or declining during the analysis period and they scored the CI items accordingly. However, it is highly unlikely that the raters would have prior knowledge regarding the trend in forecast accuracy.

Results

The Effect of PipeLine on Forecast Accuracy

Table 6 shows the relationship between using the PipeLine software for preparing a CPT and the forecast accuracy, pooled over 1995 to 2003. As expected, the mean (46 percent) and the median (27 percent) forecast error was lower when PipeLine software was used when compared to the mean (54 percent) and the median (31 percent) forecast error when it was not used. Similarly, the likelihood or the percentage of the forecast within average was higher (46 percent) when PipeLine was used, compared to when it was not used (40 percent). Regression models indicated that the observed relationship between the use of PipeLine and forecast accuracy was statistically significant¹⁵, even after accounting for variation of the forecast error due to country, client, product, and shipment accuracy (see first set of models in table 6A).

14. Robust standard errors for the median regression models were obtained using nonparametric bootstrap method (StataCorp 2003a).

15. Alpha error or p-value of less than 0.05 was considered significant for all statistical tests. Significant statistical tests are referred to as significant in the text, in brief.

Table 6. Relationship between the Use of PipeLine and Forecast Accuracy, 1995–2003 Pooled

		Forecast Error		% of Forecasts within Average	n
		Mean	Median		
PipeLine use	No	54.4	31.3	40.3	521
	Yes	45.9	27.2	46.1	529
Total		50.1	28.9	43.2	1,050

However, when the effect of a trend on the forecast accuracy was also controlled, the significant relationship between the use of PipeLine and the forecast accuracy disappeared (see second set of models in table 6A). Comparing the first set of models with the second set of models in table 6A indicated that the effect of PipeLine was collinear with the effect of trend, i.e., the use of PipeLine software explained the portion of the variation of the forecast error that was also explained by the trend effect. The finding was not surprising because the variable indicating the linear trend and the variable indicating the use of PipeLine was highly correlated; the correlation coefficient between the two variables was 0.8.

The Effect of an LMIS on Forecast Accuracy

The sample that was selected (n=207) for the analysis of the LMIS effect on forecast accuracy was first compared to the sample that was not selected (n=149) during the same forecast period (i.e., 1995, 1999, and 2000). The comparison was done to assess the bias due to the sample selection (see table 7A). The selected sample did not vary significantly by the three indicators measuring forecast accuracy. However, the distribution of the two sample were significantly different by region and client category. The distribution of the selected sample by region was 45 percent Africa, 11 percent ANE, and 44 percent LAC; while the regional distribution in the sample not selected was 54 percent Africa, 17 percent ANE, and 30 percent LAC. The client category in the selected sample was mainly MOH (73 percent) followed by other NGOs (24 percent); the SM represented a very small portion (3 percent). While less than half (46 percent) the client category in the sample not selected was for the MOH; the rest were more or less equally divided between SM and other NGO (26 and 28 percent, respectively).

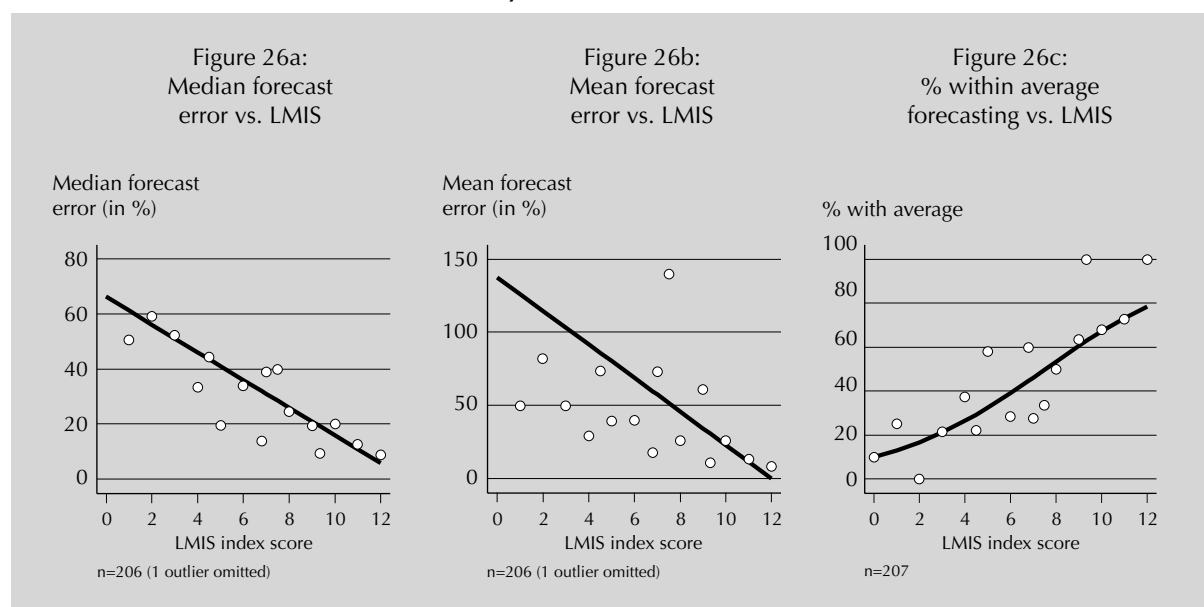
Table 7 shows the description of the forecast error and the LMIS score index in the sub-sample (i.e., n=207). The mean LMIS score in the selected sample appeared to increase from 10.7 in 1995 to 15.0 in 1999, and then decrease slightly to 13.4 in 2000. Correspondingly, the three indicators of the forecast accuracy in table 7 showed that the forecast accuracy increased between 1995 and 1999 and then decreased between 1999 and 2000, suggesting the possible relationship between forecast accuracy and LMIS score, as expected.

Table 7. Description of the Forecast Error and LMIS Index Score by Year

Forecast Year	Forecast Error		% within Average Forecasting	LMIS Score		n
	Median	Mean		Mean	Std. Dev.	
1995	36.6	73.2	32.6	10.7	5.0	86
1999	16.0	42.4	68.8	15.0	4.5	48
2000	28.6	57.3	45.2	13.4	4.1	73
Total	28.1	60.4	45.4	12.7	4.9	207

The relationship between the forecast accuracy and the LMIS score was further analyzed by plotting the three forecast accuracy indicators against the LMIS scores for 1995, 1999, and 2004 (see figures 26a, 26b, and 26c). For each level of the LMIS score, the mean, the median, and the percentage average forecasting was estimated. For example, in figure 26a each dot represents the median forecast accuracy for a given level of the LMIS score. The analysis indicates that the forecast accuracy significantly improved with the increase in the LMIS score. However, the figures 26a, 26b, and 26c are grossly confounded by the trend effect. Regression models indicated that the observed relationship between forecast accuracy and LMIS score remained significant, even after accounting for a variation of the forecast error due to trend, shipment accuracy, country, client, and product (see table 8A).

Figure 26. Scatter Plots between Forecast Accuracy and LMIS Index Score, 1995, 1999, and 2000

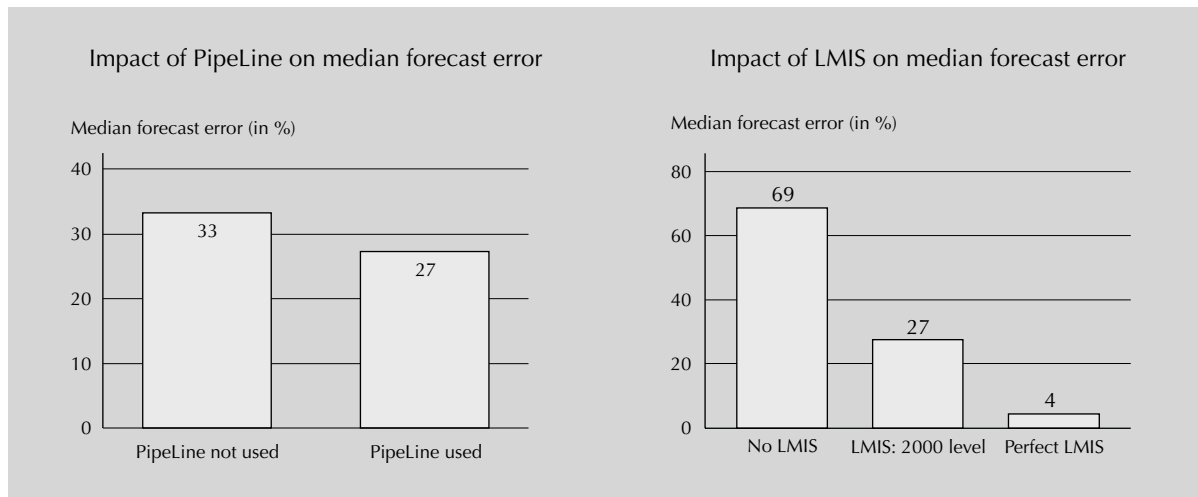


Simulation of the Effect of PipeLine and the LMIS on Forecast Accuracy

The final step of this analysis focuses on quantifying the impact of PipeLine and the LMIS on the forecast accuracy. Of the three indicators used to measure forecast accuracy, the median forecast accuracy measure was preferred because it avoided the influence of outliers¹⁶. The impact of PipeLine on the median forecast accuracy was assessed by simulating the median regression model from the first set of models in table 6A; and the impact of LMIS on the median forecast accuracy was assessed by simulating the median regression model in table 8A (see figure 27). The simulation exercise show that the use of PipeLine decreased the median forecast error by a moderate 6 percentage points, from 33 percent when PipeLine was not used to 27 percent when PipeLine was used. Since 100 percent of the CPTs in the sample are currently prepared using the PipeLine software (see figure 23), this analysis suggests that the maximum possible impact of PipeLine on forecast accuracy have already been achieved among the referenced CPT clients.

Simulation of the median regression model in table 8A show that the median forecast error was 69 percent when there was no LMIS (i.e., LMIS score=0); it was 27 percent when the LMIS functional level was at the level observed in 2000 (i.e., LMIS score = 13.4). Therefore, a 42 percentage points (69–27=42) reduction in the median forecast accuracy is attributable to the LMIS functioning at the 2000 level among the CPT clients included in the analysis. The median forecast error was only 4 percent when the functioning level of the LMIS was perfect (i.e., LMIS score = 12), indicating that the median forecast error can be further improved by another 23 percentage points (27–4=23) by improving the LMIS of the CPT clients to almost perfection (see figure 27).

Figures 27.
Impact of PipeLine and LMIS on Forecast Accuracy



16. As discussed earlier, the indicator ‘percent within average forecasting’ also avoids the influence of outliers, but it did not quantify the extent of the forecast error; therefore, it was not used for the simulation exercise.

Discussion

This study sought empirical evidence on the effect of the use of PipeLine software and the functionality of the LMIS on forecast accuracy. After minimizing the sources of bias, the study found that the use of the PipeLine software in preparing a CPT moderately improved the forecast accuracy, while the improvement in the functional level of the client's LMIS substantially improved the forecast accuracy. There is an opportunity to improve the functional level and the sustainability of the LMIS of the CPT clients further, which would lead to further improvement in the forecast accuracy for contraceptives. Contraceptive forecast accuracy may lead to contraceptive security by ensuring product availability to the family planning method users. However, further study will be required to establish that fact by looking for the empirical evidence for the relationship between forecast accuracy and product availability.

The sample for the analysis of the effect of the LMIS on forecast accuracy was not a random sample; it depended on the availability of the CI scores, and raises the question of whether the findings from the sub-sample are biased and whether the findings are applicable to all DELIVER/FPLMs CPT clients. The tool for the CI was implemented in countries where the USAID Mission Office provided financial support to carry out the assessment, which is unlikely to be influenced by the level of forecast accuracy of the CPTs done for the country. The comparison of the forecast accuracy between the sample that was selected and was not selected confirmed that the forecast error was not different between the two samples. Therefore, it can be concluded that the sample selection process did not bias the observed relationship between the LMIS and forecast accuracy. However, the selected sample represented mainly the MOH clients in the Africa and LAC region; generalizing, therefore, the findings to all the CPT clients should be done cautiously.

Another by-product of this analysis is worth mentioning: the validity of the construct measuring the functionality of the LMIS of a client using the four items in the CI tool. The relationship between the LMIS and forecast accuracy in the expected direction indicates that the construct measuring the LMIS of a client is not as invalid as might be supposed because of the criticism of the CI. Nevertheless, to minimize the measurement error, DELIVER has redesigned the tool for the CI. To evaluate the project, DELIVER is currently implementing a new CI tool, the Logistics System Assessment Tool (LSAT), in countries where it provides technical assistance (DELIVER 2002). The data from LSAT will provide another opportunity to validate the findings of this study.

Please note that as mentioned earlier, the definition of forecast accuracy used for this research is obtained from an earlier study conducted by Wilson in 1995 for comparability to that study. However, forecast accuracy is conventionally defined as the absolute difference between projected and actual consumption for a product for a given reference period, times 100, and then divided by the actual consumption. The numerators of the two definitions are the same while the denominators differ. In this paper, the denominator is the projected consumption (see page 1). As such, it is expected that the two different definitions of forecast accuracy give two different values. Nevertheless, it should be noted that the conclusions of this paper regarding improvement in forecast accuracy over time and difference in forecast accuracy by method, country, and client category would follow the same trend even if the conventional definition of forecast accuracy was used. For example, regression analysis using the conventional definition show that the median forecast accuracy has significantly improved by about 1.8 percentage points per year, from about 35 percent during 1995–1996 to about 28 percent during 2002–2003 which is similar to the trend observed in this paper (see page 7).

References

- DELIVER. 2000. *DELIVER: Contraceptives and Logistics Management, Briefing Book*. Arlington, Va.: John Snow, Inc./DELIVER.
- DELIVER. 2002. *Logistics Indicators and Monitoring and Evaluation Tools*. Arlington, Va.: John Snow, Inc./DELIVER, for the U.S. Agency for International Development.
- Family Planning Logistics Management (FPLM). 1999. *Composite Indicators for Contraceptive Logistics Management*. Arlington, Va.: FPLM, for the U.S. Agency for International Development. (Compiled by the EVALUATION project and John Snow, Inc./FPLM and the Centers for Disease Control and Prevention).
- Family Planning Logistics Management (FPLM). 2000. *The Contraceptive Forecasting Handbook for Family Planning and HIV Prevention Programs*. Arlington, Va.: FPLM/John Snow, Inc., for the U.S. Agency for International Development.
- Gelfeld, Dana. 2000. "Summary of Results: The Composite Indicators for Contraceptive Logistics Management under FPLM III." Working Paper, FPLM III. Arlington, Va.: John Snow, Inc./DELIVER, for the U.S. Agency for International Development.
- John Snow Inc. (JSI)/DELIVER. 2004. *The Logistics Handbook: A Practical Guide for Supply Chain Managers in Family Planning and Health Programs*. Arlington, Va.: JSI/DELIVER, for the U.S. Agency for International Development.
- Koenker, Roger. 2004. "Quantile regression for longitudinal data." Working Paper. Urbana-Champaign, Ill.: University of Illinois at Urbana-Champaign.
- StataCorp. 2003. *Stata Statistical Software: Release 8.0*. College Station, Tx.: Stata Corporation.
- Wilson, Edward. 1995. 1993 "Annual Accuracy Analysis of FPLM Contraceptive Forecasting." Working paper. Family Planning Logistics Management (FPLM). Arlington, Va.: John Snow, Inc./FPLM.
- Wooldridge, Jeffrey M. 2003. *Introductory Econometrics: A Modern Approach, 2nd Edition*. Mason, Ohio: Thomson, South-Western.

Appendix

Figure 1A. Format of a CPT Report

2004 Contraceptive Procurement Table					
Country:			Prepared by:		
Program:			Date prepared: December 8, 2003		
Contraceptive:					
			(CPT year)	(CPT planning year)	
	2002	2003	2004	2005	2006
1. Beginning of year stock					
2. Received/expected					
(a) Receive					
(b) Expected					
(c) Transfers/adjustments in					
3. Estimated dispensed			*	*	*
(a) Dispensed to users					
(b) Losses/transfers out					
(c) Adjustments out					
4. End of year stock (EOYS)					
5. Desired EOYS					
6. Surplus (+) or quantity needed (-)					
7. Quantity proposed					
8. Surplus (+) or shortfall (-)					

* Forecasted use.

The shaded cells are not completed.

Table 1A. Trend in the Median Forecast Accuracy by Background Characteristics, 1995–2003 (sample size in parenthesis)

Characteristic	Pooled	Year									Trend Test		Difference in Trend (p-value)	
		1995	1996	1997	1998	1999	2000	2001	2002	2003	Coef.	p-value		
Method*														0.024
Condom	28.2 (218)	39.2 (34)	36.0 (27)	32.8 (26)	21.0 (19)	24.5 (20)	22.8 (24)	18.8 (26)	34.3 (25)	51.2 (17)	-0.8	0.423		
Oral pill	26.7 (360)	30.3 (46)	40.9 (37)	31.3 (39)	41.5 (26)	26.1 (33)	26.6 (44)	24.5 (54)	24.0 (50)	22.0 (31)	-2.6	0.001		
Injectable	30.1 (208)	52.9 (19)	60.0 (21)	27.0 (22)	42.0 (18)	19.7 (20)	49.8 (27)	22.2 (28)	27.2 (33)	25.1 (20)	-3.9	0.005		
IUD	29.6 (179)	24.7 (26)	30.9 (24)	31.6 (23)	26.7 (15)	20.1 (1)	39.6 (19)	40.0 (23)	33.3 (19)	17.8 (14)	-0.7	0.564		
Implant	45.0 (85)	68.8 (9)	34.1 (7)	42.9 (9)	26.6 (8)	40.7 (7)	67.0 (12)	59.3 (14)	26.3 (12)	41.8 (7)	-0.1	0.959		
Region*														0.132
Africa	31.7 (523)	32.0 (50)	29.9 (35)	36.4 (47)	43.9 (43)	26.5 (44)	45.2 (79)	32.4 (84)	25.3 (90)	26.6 (51)	-1.7	0.008		
ANE	18.4 (128)	20.9 (17)	26.9 (16)	25.3 (17)	14.8 (18)	17.2 (20)	8.0 (10)	19.4 (18)	11.7 (8)	12.1 (4)	-1.5	0.141		
LAC	29.6 (399)	44.3 (67)	43.0 (65)	29.2 (55)	25.5 (25)	20.1 (32)	27.2 (37)	20.1 (43)	34.3 (41)	25.5 (34)	-3.2	<.001		
Client*														0.027
MOH	27.1 (634)	34.2 (86)	37.3 (70)	33.0 (76)	30.7 (51)	20.0 (65)	27.9 (69)	19.2 (81)	23.8 (81)	19.2 (55)	-2.1	<.001		
SM	26.3 (144)	30.7 (9)	13.8 (10)	20.9 (15)	25.2 (17)	33.3 (19)	30.8 (17)	18.6 (23)	21.6 (20)	29.1 (14)	-0.5	0.706		
Other NGO	39.7 (272)	39.7 (39)	32.5 (36)	25.3 (28)	33.4 (18)	33.7 (12)	48.7 (40)	50.0 (41)	38.9 (38)	38.2 (20)	0.4	0.622		
No. of donor														0.539
Single	29.7 (757)	37.1 (93)	33.8 (75)	28.2 (77)	27.3 (62)	22.2 (81)	40.0 (91)	27.0 (105)	30.2 (103)	27.5 (70)	-1.7	0.014		
Multiple	25.9 (248)	23.3 (36)	33.5 (28)	28.6 (25)	36.9 (22)	24.4 (15)	28.3 (32)	21.5 (36)	20.9 (35)	14.6 (19)	-2.3	0.002		
TOTAL	28.9 (1,050)	34.8 (134)	34.0 (116)	29.2 (119)	29.5 (86)	23.3 (96)	34.2 (126)	25.9 (145)	27.2 (139)	25.5 (89)	-1.9	0.001		

Notes for the table are given in the Appendix Endnotes.

Table 2A. Trend in the Mean Forecast Accuracy by Background Characteristics, 1995–2003
(sample size in parenthesis)

Characteristic	Pooled	Year									Trend Test		Difference in Trend (p-value)
		1995	1996	1997	1998	1999	2000	2001	2002	2003	Coef.	p-value	
Method*													0.002
Condom	38.5 (218)	48.7 (34)	43.3 (27)	38.6 (26)	32.3 (19)	32.1 (20)	40.0 (24)	24.3 (26)	37.1 (25)	46.2 (17)	-1.4	0.647	
Oral pill	42.8 (360)	48.6 (46)	40.7 (37)	92.3 (39)	40.0 (26)	40.9 (33)	39.1 (44)	30.2 (54)	29.0 (50)	27.8 (31)	-3.4	0.147	
Injectable	67.2 (208)	157.2 (19)	83.8 (21)	41.9 (22)	126.1 (18)	79.3 (20)	69.0 (27)	35.8 (28)	32.0 (33)	26.4 (20)	-15.7	0.002	
IUD	44.0 (179)	33.2 (26)	36.0 (24)	37.1 (23)	37.8 (15)	30.5 (16)	40.1 (19)	72.8 (23)	69.7 (19)	34.1 (14)	2.2	0.332	
Implant	82.5 (85)	65.8 (9)	69.0 (7)	64.0 (9)	34.0 (8)	39.6 (7)	190.8 (12)	125.9 (14)	37.4 (12)	44.4 (7)	-4.0	0.566	
Region													0.009
Africa	49.0 (523)	41.2 (50)	31.5 (35)	49.9 (47)	45.7 (43)	46.5 (44)	69.9 (79)	61.1 (84)	40.6 (90)	35.3 (51)	-0.7	0.701	
ANE	24.4 (128)	25.0 (17)	32.0 (16)	30.1 (17)	18.6 (18)	26.6 (20)	16.4 (10)	24.8 (18)	19.5 (8)	11.7 (4)	-1.2	0.276	
LAC	59.9 (399)	87.2 (67)	64.1 (65)	74.5 (55)	98.6 (25)	55.0 (32)	57.7 (37)	26.1 (43)	34.2 (41)	32.9 (34)	-9.3	<.001	
Client*													0.598
MOH	44.5 (634)	72.4 (86)	58.5 (70)	42.4 (76)	40.2 (51)	25.1 (65)	48.4 (69)	36.2 (81)	38.0 (81)	29.3 (55)	-4.8	0.007	
SM	40.1 (144)	36.4 (9)	25.4 (10)	58.5 (15)	39.1 (17)	51.8 (19)	45.7 (17)	34.3 (23)	27.0 (20)	39.8 (14)	-1.4	0.757	
Other NGO	68.7	45.5	39.8	101.9	113.9	144.0	87.0	72.8	41.8	39.8	-2.0	0.458	
No. of donor													0.245
Single	47.0 (757)	51.7 (93)	44.7 (75)	64.3 (77)	59.8 (62)	44.9 (81)	59.0 (91)	34.6 (105)	35.0 (103)	36.0 (70)	-3.1	0.039	
Multiple	53.7 (248)	90.9 (36)	63.0 (28)	35.0 (25)	44.6 (22)	47.0 (15)	64.6 (32)	59.8 (36)	30.0 (35)	23.5 (19)	-6.6	0.009	
TOTAL	50.1 (1,050)	62.2 (134)	49.8 (116)	58.4 (119)	55.4 (86)	45.2 (96)	60.3 (126)	46.2 (145)	37.5 (139)	33.3 (89)	-4.4	0.002	

Notes for the table are given in the Appendix Endnotes.

Table 3A. Trend in the Median Forecast Error by Country, All Methods, 1995–2003 (sample size in parenthesis)

Country	Year									Trend Effect	
	1995	1996	1997	1998	1999	2000	2001	2002	2003	Coef.	p-value
Bangladesh	na	na	na	na	29.0 (2)	55.0 (1)	57.3 (2)	38.5 (3)	na	-0.7	0.978
Bolivia	32.7 (13)	13.5 (19)	27.2 (12)	714.0 (2)	1013.3 (1)	47.4 (7)	36.2 (6)	8.3 (3)	38.3 (3)	0.7	0.895
Burkina Faso	na	na	na	na	na	63.1 (6)	59.3 (6)	44.5 (6)	na	-10.0	0.260
Cameroon	na	na	na	na	na	68.8 (14)	46.6 (16)	19.3 (15)	22.9 (8)	-12.9	0.010
Egypt	22.4 (4)	44.7 (4)	49.2 (4)	34.1 (4)	27.4 (4)	7.9 (4)	16.7 (2)	na	na	-3.6	0.456
El Salvador	51.9 (14)	60.0 (13)	na	na	na	na	27.6 (5)	na	19.3 (12)	-5.0	0.005
Ghana	25.3 (10)	29.9 (11)	45.9 (12)	43.0 (18)	40.5 (14)	40.0 (17)	44.6 (18)	30.2 (19)	30.0 (11)	-1.3	0.267
Guatemala	39.7 (13)	36.2 (13)	36.0 (13)	26.1 (14)	13.1 (14)	22.4 (19)	22.1 (14)	38.1 (12)	29.3 (19)	-1.8	0.174
Haiti	269.0 (6)	70.9 (6)	18.1 (7)	na	14.7 (6)	na	24.2 (6)	16.6 (6)	na	-6.8	0.061
Malawi	59.3 (5)	29.5 (5)	66.6 (4)	56.3 (6)	8.7 (4)	24.6 (6)	16.9 (6)	17.6 (6)	39.7 (6)	-2.6	0.148
Mali	30.7 (13)	12.5 (5)	71.2 (6)	38.4 (4)	66.4 (3)	35.7 (6)	50.0 (2)	39.0 (3)	50.0 (9)	2.3	0.218
Nepal	19.4 (10)	34.3 (8)	21.4 (9)	14.8 (10)	20.4 (10)	4.8 (5)	19.4 (10)	10.9 (5)	na	-1.7	0.285
Nicaragua	40.0 (7)	na	34.7 (9)	18.5 (9)	26.1 (6)	27.9 (11)	14.2 (7)	33.3 (13)	na	-0.7	0.804
Peru	47.4 (14)	37.4 (14)	29.4 (14)	na	28.1 (5)	na	16.4 (5)	41.2 (7)	na	-2.1	0.630
Philippines	20.9 (3)	17.5 (4)	12.8 (4)	8.6 (4)	9.1 (4)	na	5.1 (4)	na	12.1 (4)	-0.5	0.650
Tanzania	33.3 (7)	38.2 (6)	22.3 (6)	na	18.2 (8)	na	18.7 (9)	31.7 (9)	11.5 (9)	-2.9	0.029
Togo	na	na	na	na	na	45.2 (15)	27.4 (13)	23.3 (17)	26.5 (8)	-1.1	0.792
Uganda	9.5 (8)	na	51.2 (11)	54.9 (9)	50.3 (9)	28.9 (9)	35.9 (9)	19.9 (10)	na	0.6	0.831
Zimbabwe	15.4 (7)	21.4 (8)	11.4 (8)	17.7 (6)	25.4 (6)	17.1 (6)	25.9 (5)	20.9 (5)	na	0.6	0.555

Notes for the table are given in the Appendix Endnotes.

Table 4A. Trend in the Percentage of the Forecast for All Methods That Overestimated or Underestimated Actual Consumption by More Than 25% and the Percentage of the Forecast Within $\pm 25\%$ of the Actual Consumption, According to Background Characteristics, 1995–2003

Characteristic		Year										Trend Effect		Difference in Trend (p-value)	
		Pooled	1995	1996	1997	1998	1999	2000	2001	2002	2003	Coef.	p-value		
Method*															0.149
CondomUnder	21.1	29.4	29.6	26.9	15.8	15.0	25.0	11.5	12.0	17.7	-0.15	0.063			
	Average	45.4	35.3	33.3	34.6	52.6	55.0	54.2	73.1	48.0	23.5	0.11	0.093		
	Over	33.5	35.3	37.0	38.5	31.6	30.0	20.8	15.4	40.0	58.8	-0.01	0.880		
Pill	Under	19.2	30.4	21.6	18.0	19.2	21.2	13.6	14.8	20.0	12.9	-0.15	0.030		
	Average	46.9	39.1	40.5	46.2	38.5	48.5	50.0	50.0	50.0	58.1	0.17	0.003		
	Over	33.9	30.4	37.8	35.9	42.3	30.3	36.4	35.2	30.0	29.0	-0.08	0.181		
Injectable	Under	29.8	52.6	47.6	36.4	38.9	20.0	37.0	17.9	15.2	15.0	-0.20	0.007		
	Average	38.9	15.8	28.6	45.5	33.3	55.0	22.2	50.0	45.5	50.0	0.16	0.024		
	Over	31.3	31.6	23.8	18.2	27.8	25.0	40.7	32.1	39.4	35.0	0.04	0.631		
IUD	Under	10.1	11.5	16.7	13.0	6.7	6.3	0.0	13.0	15.8	0.0	-0.19	0.100		
	Average	42.5	50.0	37.5	39.1	40.0	62.5	31.6	30.4	36.8	64.3	0.07	0.335		
	Over	47.5	38.5	45.8	47.8	53.3	31.3	68.4	56.5	47.4	35.7	0.01	0.893		
Implant	Under	34.1	22.2	28.6	22.2	12.5	42.9	50.0	50.0	25.0	42.9	0.11	0.347		
	Average	34.1	22.2	42.9	44.4	50.0	42.9	16.7	21.4	50.0	28.6	-0.04	0.742		
	Over	31.8	55.6	28.6	33.3	37.5	14.3	33.3	28.6	25.0	28.6	-0.10	0.461		
Region*															0.047
AFR	Under	20.1	32.0	22.9	14.9	16.3	15.9	20.3	17.9	21.1	19.6	-0.04	0.490		
	Average	40.2	40.0	42.9	40.4	27.9	45.5	29.1	39.3	50.0	45.1	0.06	0.196		
	Over	39.8	28.0	34.3	44.7	55.8	38.6	50.6	42.9	28.9	35.3	-0.03	0.478		
ANE	Under	18.0	17.7	25.0	17.7	11.1	25.0	10.0	16.7	25.0	0.0	-0.19	0.199		
	Average	60.9	58.8	37.5	47.1	66.7	60.0	80.0	72.2	62.5	100.0	0.32	0.008		
	Over	21.1	23.5	37.5	35.3	22.2	15.0	10.0	11.1	12.5	0.0	-0.26	0.049		
LAC	Under	24.1	29.9	30.8	30.9	32.0	18.8	29.7	18.6	7.3	8.8	-0.27	<.001		
	Average	41.6	26.9	32.3	41.8	48.0	59.4	48.7	55.8	36.6	47.1	0.15	0.001		
	Over	34.3	43.3	36.9	27.3	20.0	21.9	21.6	25.6	56.1	44.1	0.03	0.617		
Client*															<.001
MOH	Under	21.9	33.7	40.0	23.7	21.6	18.5	15.9	14.8	12.4	14.6	-0.21	<.001		
	Average	46.2	38.4	30.0	38.2	39.2	60.0	46.4	56.8	50.6	58.2	0.18	<.001		
	Over	31.9	27.9	30.0	38.2	39.2	21.5	37.7	28.4	37.0	27.3	-0.04	0.360		
SM	Under	23.6	33.3	0.0	26.7	11.8	21.1	35.3	34.8	20.0	21.4	0.09	0.379		
	Average	45.8	22.2	60.0	53.3	47.1	36.8	29.4	56.5	60.0	35.7	-0.13	0.183		
	Over	30.6	44.4	40.0	20.0	41.2	42.1	35.3	8.7	20.0	42.9	0.06	0.598		
Other NGO	Under	18.8	18.0	11.1	17.9	22.2	16.7	27.5	14.6	26.3	10.0	-0.04	0.681		
	Average	34.9	33.3	41.7	46.4	44.4	41.7	30.0	26.8	31.6	30.0	0.05	0.493		
	Over	46.3	48.7	47.2	35.7	33.3	41.7	42.5	58.5	42.1	60.0	-0.03	0.718		
No. of donors*															0.583
Single	Under	20.0	25.8	22.7	26.0	19.4	18.5	20.9	17.1	15.5	14.3	-0.14	0.001		
	Average	41.9	30.1	37.3	42.9	45.2	50.6	38.5	45.7	44.7	42.9	0.11	<.001		
	Over	38.2	44.1	40.0	31.2	35.5	30.9	40.7	37.1	39.8	42.9	-0.02	0.510		
Multiple	Under	27.0	41.7	42.9	24.0	22.7	20.0	28.1	19.4	20.0	15.8	-0.13	0.012		
	Average	48.8	50.0	35.7	44.0	31.8	66.7	40.6	55.6	54.3	68.4	0.12	0.006		
	Over	24.2	8.3	21.4	32.0	45.5	13.3	31.3	25.0	25.7	15.8	-0.03	0.497		
TOTAL	Under	21.3	29.1	27.6	22.7	19.8	18.8	22.2	17.9	17.3	14.6	-0.14	<.001		
Average	43.2	35.8	36.2	42.0	41.9	53.1	38.9	48.3	46.8	48.3	0.12	<.001			
Over	35.4	35.1	36.2	35.3	38.4	28.1	38.9	33.8	33.8	36.0	37.1	-0.02	0.462		

Notes for the table are given in the Appendix Endnotes.

Table 5A. Trend in the Sum of Projected and Actual Consumption (in 1,000s) by Background Characteristics and the Percentage Difference Between Them, by Background Characteristics, 1995–2003

Background Characteristics		Year								
		1995	1996	1997	1998	1999	2000	2001	2002	2003
Method										
Condom	Projected	165,641	143,821	176,930	157,123	351,520	148,243	186,232	243,419	128,927
	Actual	161,429	148,051	158,633	128,182	294,038	129,553	176,219	234,712	115,566
	Pct. diff.	2.5	-2.9	10.3	18.4	16.4	12.6	5.4	3.6	10.4
Oral pill	Projected	37,393	36,636	37,547	30,008	43,761	22,730	58,195	57,913	28,829
	Actual	33,571	32,615	32,828	27,400	44,224	18,996	63,789	63,145	25,490
	Pct. diff.	10.2	11.0	12.6	8.7	-1.1	16.4	-9.6	-9.0	11.6
Injectable	Projected	3,012	5,171	9,458	7,890	11,660	8,039	15,229	15,277	7,613
	Actual	4,065	6,317	8,125	6,779	11,523	8,315	14,363	12,376	7,563
	Pct. diff.	-35.0	-22.2	14.1	14.1	1.2	-3.4	5.7	19.0	0.7
IUD	Projected	1,391	1,387	1,394	1,452	1,749	1,444	1,669	168	141
	Actual	1,349	1,478	1,735	1,416	1,429	1,236	1,702	115	123
	Pct. diff.	3.0	-6.6	-24.4	2.5	18.3	14.4	-2.0	31.2	12.8
Implant	Projected	30	24	34	31	44	54	51	65	33
	Actual	14	18	42	24	42	57	60	72	32
	Pct. diff.	53.3	22.6	-21.9	22.1	4.8	-5.4	-18.0	-9.5	2.7
Region										
AFR	Projected	111,431	86,417	124,458	124,878	133,387	139,216	133,181	209,252	124,898
	Actual	105,385	97,205	112,306	97,897	122,195	118,496	128,400	210,786	117,993
	Pct. diff.	5.4	-12.5	9.8	21.6	8.4	14.9	3.6	-0.7	5.5
ANE	Projected	59,524	63,338	60,330	57,700	236,709	22,889	71,080	39,688	21,272
	Actual	53,212	48,700	48,198	51,896	189,708	23,775	74,307	47,595	17,530
	Pct. diff.	10.6	23.1	20.1	10.1	19.9	-3.9	-4.5	-19.9	17.6
LAC	Projected	36,512	37,284	40,575	13,926	38,638	18,404	57,114	67,901	19,373
	Actual	41,831	42,573	40,858	14,008	39,352	15,885	53,426	52,039	13,251
	Pct. diff.	-14.6	-14.2	-0.7	-0.6	-1.8	13.7	6.5	23.4	31.6
Client										
MOH	Projected	147,510	134,546	170,271	138,615	158,982	112,207	146,727	183,044	80,964
	Actual	153,275	139,701	142,275	120,366	174,301	93,271	134,802	171,981	79,227
	Pct. diff.	-3.9	-3.8	16.4	13.2	-9.6	16.9	8.1	6.0	2.1
SM	Projected	21,273	19,716	32,629	44,163	239,462	50,234	92,348	113,098	72,524
	Actual	21,044	15,943	34,646	31,439	168,712	49,360	106,217	120,809	63,181
	Pct. diff.	1.1	19.1	-6.2	28.8	29.5	1.7	-15.0	-6.8	12.9
Other NGO	Projected	38,684	32,776	22,463	13,725	10,290	18,068	22,300	20,699	12,055
	Actual	26,109	32,835	24,441	11,996	8,243	15,525	15,115	17,629	6,364
	Pct. diff.	32.5	-0.2	-8.8	12.6	19.9	14.1	32.2	14.8	47.2
No. of donors										
Single	Projected	156,546	128,556	173,771	166,471	231,977	69,493	157,676	183,567	73,809
	Actual	142,338	137,605	168,802	144,918	215,828	63,071	160,315	191,250	63,583
	Pct. diff.	9.1	-7.0	2.9	12.9	7.0	9.2	-1.7	-4.2	13.9
Multiple	Projected	49,499	57,889	48,782	30,030	176,757	111,003	103,255	133,274	91,734
	Actual	57,017	50,475	32,223	18,881	135,427	95,074	95,673	119,169	85,190
	Pct. diff.	-15.2	12.8	33.9	37.1	23.4	14.3	7.3	10.6	7.1
Total	Projected	207,467	187,038	225,363	196,503	408,734	180,509	261,375	316,841	165,543
	Actual	200,429	188,479	201,362	163,800	351,256	158,156	256,133	310,420	148,773
	Pct. diff.	3.4	-0.8	10.6	16.6	14.1	12.4	2.0	2.0	10.1

Table 6A. Regression Models Predicting the Effect of PipeLine on Forecast Accuracy (n=1,050)

Independent Variable	Median Forecast Accuracy		Mean Forecast Accuracy		Likelihood of Average Forecasting	
	coef. (SE)	p-value	coef. (SE)	p-value	coef. (SE)	p-value
First set of models						
Shipment accuracy adequate (not)	-11.41 (2.36)	<.001	1.03 (7.86)	0.896	0.62 (0.18)	<.001
PipeLine used (not used)	-6.02 (2.21)	0.006	-17.71 (7.73)	0.022	0.64 (0.17)	<.001
Constant	7.59(23.22)	0.744	58.36 (7.44)	<.001	na	
Second set of models						
Shipment accuracy adequate (not)	-9.12 (2.53)	<.001	1.58 (7.85)	0.841	0.62 (0.18)	<.001
PipeLine used (not used)	3.32 (4.60)	0.471	2.58(12.80)	0.841	0.22 (0.28)	0.429
Trend	-2.56 (0.61)	<.001	-4.78 (2.41)	0.047	0.10 (0.05)	0.058
Constant	11.38 (26.03)	0.662	66.81 (8.56)	<.001		

Notes for the table are given in the Appendix Endnotes.

Table 7A. Comparison of the Characteristics of the Sample That Was Selected for the Analysis of LMIS Effect on Forecast Accuracy and the Sample That Was Not Selected (1995, 1999, and 2000 pooled)

Characteristics	Not Selected	Selected	Total
Percentage distribution of the samples			
Region*			
Africa	53.7	44.9	48.6
Asia & the Near East	16.8	10.6	13.2
Latin America & the Caribbean	29.5	44.4	38.0
Method			
Condom	22.2	21.7	21.9
Oral pill	35.6	33.8	34.6
Injectable	16.8	19.8	18.5
IUD	15.4	18.4	17.1
Implant	10.1	6.3	7.9
Client*			
Ministry of Health	46.3	73.0	61.8
Social marketing	26.2	2.9	12.6
Other NGOs	27.5	24.2	25.6
Descriptive statistics			
Forecast accuracy			
Mean	33.3	28.1	29.6
Median	52.1	60.4	56.9
Percentage average	36.2	45.4	41.6
Sample size	149	207	356

* Significant ($p < .05$) variation between the two sample.

Table 8A. Regression Models Predicting the Effect of LMIS on Forecast Accuracy (n=207)

Independent Variable	Median Forecast Accuracy			Mean Forecast Accuracy			Likelihood of Average Forecasting		
	<i>coef.</i>	<i>(SE)</i>	<i>p-value</i>	<i>coef.</i>	<i>(SE)</i>	<i>p-value</i>	<i>coef.</i>	<i>(SE)</i>	<i>p-value</i>
Shipment accuracy adequate (not)	1.46	(7.96)	0.854	24.99	(24.19)	0.305	-0.22	(0.59)	0.707
Trend	-1.08	(2.10)	0.608	4.60	(6.23)	0.462	0.045	(0.13)	0.722
LMIS score	-5.37	(1.73)	0.002	-12.97	(6.01)	0.034	0.29	(0.15)	0.056
Constant	33.13	(276.86)	0.905	118.27	(34.30)	0.001	na		

Notes for the table are given in the Appendix Endnotes.

Appendix Endnotes

Table 1A.

NOTE: Two types of analysis were performed for table 1A, pooled and trend analysis: (1) Pooled analysis: The forecast accuracy measures for all method during the period 1995 to 2003 were pooled to assess their association with the background characteristics (i.e., the independent variables). Median regression model with robust standard errors (i.e., bootstrap method) (StataCorp 2003) was used to assess the association. The models accounted for the repeated measures (of countries, client, and product) and survey year. The independent variables that were significantly related (i.e., the probability that the observed relationship is due to chance is less than 5 percent, also referred to as $p < .05$) to the median forecast error is marked with an asterisk (*) in table 1A. Method, region, and client categories were significantly ($p < .05$) associated with the forecast error. (2) Trend analysis: This was conducted to assess significant trends in the median forecast error for each of the categories of the independent variables and the variation of the trend within the categories of an independent variable. Median regression models with robust standard errors were used for the purpose. The models accounted for the repeated measures. The coefficients of the trend effect and its significant level (i.e., the p-value) are reported in table 1A. Negative coefficient indicates a declining trend, while a positive coefficient indicates an increasing trend in the forecast error. Having a p-value for the trend effects less than 0.05 and indicates a significant trend effect. For example, the coefficient and the p-value of the trend effect for pill in table 1A indicates that the median forecast error is significantly (because the $p = .001$ which is less than 0.05) declining by an average of 2.6 percentage points per year from 1995 to 2003. The p-value of the test that determined differences in the trend of the forecast error within the categories of an independent variable is also reported in the last column of table 1A. For example, the difference in trend p-value (0.027) for client type indicated that the trend of the median forecast error between at least two client categories was significantly ($p < .05$) different from each other. Please contact the author for the tables of the models used in this analysis.

Table 2A.

NOTE: The analysis in table 2A repeats the analysis in table 1A, using the mean forecast error instead of the median; uses fixed-effects linear regression models (StataCorp 2003) instead of the median regression models. The background characteristics that were significantly ($p < .05$) related with the pooled forecast accuracy are marked with an asterisk (*). The interpretation of the coefficients and p-values in table 2A are analogous to those in table 1A. For example, the coefficient and the p-value of the trend effect for injectable in table 2A indicates that the median forecast error is significant ($p = .002$), declining by an average of 15.7 percentage points per year from 1995 to 2003. Please contact the author for the tables of the models used in this analysis.

Table 3A.

NOTE: An 'na' indicates missing data.

The trend effect of the median forecast accuracy for each country is assessed using median regression models that were controlled for method and client category. The coefficient of the trend effect and its p-value are reported and interpreted similarly to those in tables 1A and 2A. For example, the trend in the median forecast error in Tanzania is significantly ($p = 0.029$) declining by 2.9 percentage points per year.

Table 4A.

NOTE: The analysis in table 4A is similar to those in table 1A and 2A. The background characteristics that were significantly ($p < .05$) related to the pooled outcome is marked with an asterisk (*). Pooled and trend analysis was conducted. Because the outcome variable in this table is categorical (under, average, and over), multinomial logit or fixed-effects logit models (StataCorp 2003) are used as appropriate instead of mean or median regression. The interpretation of the coefficient of the trend effect of the logit model, in terms of the direction, is similar to the models in tables 1A, 2A, and 3A. A positive coefficient of the trend effect indicates increasing trend and a negative coefficient indicates decreasing trend. However, exponential transformation of the coefficient is required to obtain the average change of the outcome over time. For the sake of simplicity, only the direction of the coefficients is interpreted in this analysis to indicate increasing or decreasing trends. For example, the trend in the probability of the projected use to under-estimate actual use by more than 25 percent for all methods

decreased ($p < .001$) from 1995 to 2003, while the probability that the projected use was within ± 25 percent of the actual use (referred to as the average in table 4A) increased ($p < .001$) during the same period. The p-value of the test whether the trend an overestimate, average, and underestimate varied by the categories of a background characteristic is also reported. The p-values in the difference in trends indicate that the trend in an overestimate, average, and underestimate varied by region ($p = 0.047$) and client category ($p < .001$). Please contact the author for the tables of the models used in this analysis.

Table 6A.

NOTE: An 'na' indicates not applicable. Each set of models in table 6A includes a median regression model predicting median forecast error, a fixed-effects ordinary least square (OLS) model predicting the mean forecast error, and a fixed-effects logit model predicting the likelihood of the forecast being within ± 25 percent of the actual consumption. The fixed-effects OLS and logit models held the effects of country, client, and product constant, over time. While the median regression models controlled for country, client, and product by including dummy variables, the coefficients are not shown. The main regressors for the first set of models are shipment accuracy and Pipeline use. The second set of models adds trend to the regressors of the first set of models.

For the OLS and the median regression models, the coefficient for the shipment accuracy indicates the difference between the forecast error when there is adequate shipment accuracy and when there is not adequate, i.e., the reference category for adequate shipment accuracy effect is not adequate, indicated by 'not' in parenthesis. A significant ($p < .05$) and negative coefficient for shipment accuracy adequate effect in the OLS and median regression models indicates that the forecast error is lower when shipment is adequate compared to when it is not adequate. The interpretation of the coefficient of shipment accuracy adequate in the logit models is the opposite, a positive coefficient indicates better forecast accuracy when shipment accuracy is adequate compared to when the shipment accuracy is not adequate. Similarly, the reference category for Pipeline used is not used, and the interpretation of its coefficient is similar to that of the shipment accuracy. The trend is a linear term in all models and its coefficient indicates the incremental effect of a one-year advancement in time on forecast accuracy. For the OLS and the median regression models, a significant ($p < .05$) and negative coefficient of the trend effect indicates a decrease in forecast error over time, while the interpretation of the coefficient of the trend effect is the opposite for the logit models.

The effect of Pipeline was significant in all three of the first set of models, indicating the use of Pipeline improved the forecast accuracy. However, after controlling for trend effect, the effect of Pipeline being used became not significant (and inconsistent) in all second set of models, indicating that the effect of Pipeline was collinear with the effect of trend.

Table 8A.

NOTE: An 'na' indicates not applicable. As in table 6A, the median forecast accuracy was analyzed using a median regression model; the mean forecast accuracy was analyzed using a fixed-effects OLS model; and the average forecast accuracy was analyzed using a fixed-effects logit model. The fixed-effects OLS and logit models held the effects of country, client, and product constant over time. While in the median regression models controlled for country, client, and product by including dummy variables, the coefficients of which are not shown. The main regressors for the models above are shipment accuracy, trend, and LMIS index score. The variable indicating Pipeline use was not included in the models to avoid inconsistency problems due to its collinearity with the trend effect observed in table 6A. The interpretation of the coefficients of shipment accuracy and trend are similar as it was in table 6A. The LMIS score is a linear term in all the models, and its coefficient indicates the incremental effect of one unit increase in the LMIS score on the forecast accuracy indicators. For the OLS and the median regression models a significant ($p < .05$) and negative coefficient of the LMIS score indicates that the increase in LMIS score decreases forecast error. The interpretation of the coefficient of the LMIS score is the opposite for the logit models; a positive coefficient indicates that the increase in LMIS score improves the forecast accuracy. The effect of LMIS score was significant in all three models, indicating that a comparatively better LMIS system of a client is associated with comparatively better forecast accuracy.