National Drug Stockout Risks and the Global Fund Disbursement Process for Procurement

Online Supplement

A Model Input Data

A.1 Procurement Lead Times (PLT)

We analyzed all 3027 procurement orders initiated by principal recipients in 53 African countries between January 2002 and March 2012 and recorded in the Price and Quality Recording (PQR database). The PQR database relies on the grant recipient self-reporting the order placement date as the date when the supplier accepts the order for immediate processing.

We used the PQR database to estimate the Procurement Lead Time (PLT) in our model. For the purposes of our analysis, we define PLT as the time from the ordering date by which the supplier can deliver when it is the principal recipient's intention to receive the order as soon as possible. However, an occasional practice by grant recipients is to place staggered or pre-planned orders involving the upfront communication of a schedule of several future target delivery dates; this affects 496 orders out of 3027 in the dataset. Including such observations in our PLT estimates would lead to over-estimating the procurement lead time, as for the second and subsequent requested delivery dates in these orders the lead time is presumably much higher than what the supplier could achieve (some of these orders are scheduled more than a year in advance for example). This would, in turn, result in over-estimating stockout levels. Therefore, we excluded those 496 pre-planned orders as well as 65 other outlying orders with a z-score for the lead time greater than 2.5 (these had PLTs greater than one year and so likely were also pre-planned). Consequently, the results reported in Table 3 could in fact be under-estimating stockout levels. When the 496 pre-planned and 65 outlying orders described above are included in the estimation of PLTs, the stockout level estimates reported in Table 3 increase by up to 0.9 percentage points.

We used multiple regression model specifications to identify drivers of PLT variability (see Table A1) and group the orders into 40 categories characterised by product type, region (East, North, South and West & Central Africa) and whether the receiving country was landlocked. Additional

explanatory variables were not included due to the limited sample size of each category. Categories differing on one dimension only were merged whenever the pairwise Kolmogorov-Smirnov test for identical distributions had a p-value larger than 0.5. Distributional forecasts of PLT in each category were constructed with non-parametric kernel density estimation (Greene 2012). Forecast accuracy for each category was assessed with 1000-time repeated random sub-sampling validation involving 80% randomly selected data for in-sample density estimation and comparison with remaining out-of-sample data using the Kolmogorov-Smirnov test (Arlot and Celisse 2010). A maximum of 67‰ repetitions were rejected at the 5% significance level across all categories (mean 46‰).

Multiple regression specifications (1)-(3) show the significance of drug type and African region when predicting procurement lead times (PLTs) as well as exactly one of the following variables: whether the country is landlocked or not; whether the procurement order was through voluntary pooled procurement (VPP) or not; whether there was a pre-paid component of the order. The results show that PLTs increase on average whenever any of these three characteristics are met. Specification (1) was selected for the model based on the highest R^2. Specifications (4) and (5) show that procuring through VPP and pre-payment leads to higher average PLTs, even after controlling for the geographic position of the receiving country. These result suggest that while the VPP initiative may have been successful at reducing purchasing price, it does not seem to have reduced procurement lead times, perhaps due to the additional delays associated with pooling procurement orders for the same products across different principal recipients. Furthermore, linear monthly and yearly trends were not significant in any of the models and were thus omitted - see Table A1.

The small number of variables in these models limits their R^2 to 11-14%. While adding more explanatory variables (such as country or principal recipient fixed effects) would increase their R^2 and fit (i.e., decrease their fitting error), because of the relatively small amount of data available this would negatively impact their out-of-sample prediction error which is a key criteria for the present study.

Table A2 shows summary statistics of historical PLTs for the final configuration of categories. They show substantial variations of procurement lead times across product types m (HIV prevention and treatment products take 46 days less and 18 days more than malaria prevention products on average, respectively) and geographic conditions $region(j)$ and $land(j)$ (deliveries to East Africa take 29 more days, on average, than to West & Central Africa, and landlocked countries take on

Table A1: Multiple regression model specifications for procurement lead times.

PLT Regression Specification	(1)	(2)	(3)	(4)	(5)
Drug Type					
Baseline: Malaria Prevention					
Anti-Malaria	-9.41	-6.57	-8.99	-9.17	-6.9
	(6.73)	(6.95)	(6.86)	(6.72)	(6.79)
Anti-Retroviral	18.02	18.46	17.46	19	19.81
	(7.80)**	(8.04)**	(7.98)**	(7.81)**	(7.87)**
Anti-TB	-26.96	-20.78	-25.55	-26.6	-22.1
	(6.05)***	(6.37)***	(6.24)***	(6.05)***	(6.19)***
HIV Prevention	-46.61	-42.76	-46.41	-46.58	-43.16
	(6.40)***	(6.63)***	(6.56)***	(6.40)***	(6.47)***
African region					
Baseline: West & Central					
North	-1.32	-1.54	-1.48	-1.00	-1.1.0
	(4.10)	(4.11)	(4.13)	(4.11)	(4.08)
East	28.77	31.78	32.00	29.83	29.69
	(3.92)***	(3.90)***	(3.93)***	(3.94)***	(3.91)***
South	-13.57	-12.55	-13.13	-12.67	-12.18
	(3.49)***	(3.56)***	(3.57)***	(3.50)***	(3.50)***
Landlocked	19.97			19.91	18.99
(1=yes; 0=no)	(3.01)***			(3.09)***	(3.11)***
VPP			13.26	12.89	
(1=yes; 0=no)			(6.03)**	(6.14)**	
Pre-paid		16.45			15.39
(1=yes; 0=no)		(2.99)***			(3.02)***
Observations	3,027	3,027	3,027	3,027	3,027
R-squared	0.131	0.119	0.132	0.132	0.137

Notes: Multiple regressions of PLT on type of drug, African region and: (1) geographic position; (2) whether the procurement order was pre-paid; (3) whether the order was procured through VPP; (4) geographic positiong and whether the order was procured through VPP; (5) geographic position and whether the procurement order was pre-paid. Robust standard errors in parentheses; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A2: Summary statistics of procurement lead times (PLT) in days

Number of observations, mean and (5%, 95%) quantiles of PLT in each category; categories were merged whenever the Kolmogorov-Smirnov test of common underlying distribution had a p-value larger than 0.5; merged categories are identified by either an omitted cell border, †, or ‡.

African region *region*(j)	Landlocked (Y/N) *land*(j)	Health Product Type m														
		Anti-Malarial			Anti-Retroviral			Anti-Tuberculosis			HIV Prevention			Malaria Prevention		
		n	Mean	Quantiles	n	Mean	Quantiles	n	Mean	Quantiles	n	Mean	Quantiles	n	Mean	Quantiles
East	No	222	162	(31, 347)	222	162	(31, 347)	38	231	(112, 289)	61	100	(14,236)	273	133	(31, 241)
	Yes	87	128	(35, 266)	29	231	(112, 364)	29	231	(112, 364)	60	73	(13, 151)			
South	Yes				181	76	(18, 229)	†33	122	(52, 420)				166	98	(18, 243)
North	Yes	21	103	(29, 266)	11	160	(64, 234)	11	160	(64, 234)	86	74	(17, 179)			
	No	‡62	107	(23, 280)	69	122	(30, 263)	‡62	107	(23, 280)				184	90	(11, 241)
South	No	53	146	(13, 297)				†33	122	(52, 420)	110	54	(6, 151)	550	66	(8, 168)
West & Central	No	233	115	(13, 300)	233	115	(13, 300)	233	115	(13, 300)	115	84	(13, 217)	218	84	(9,243)
	Yes	36	165	(5, 351)	18	159	(76, 206)	18	159	(76, 206)	25	93	(14, 228)	100	150	(27, 331)

Table A3: Distributional forecasts for procurement lead times: out-of-sample prediction accuracy

African region	Landlocked (Y/N)	Type of Drug														
		Anti-Malarial			Anti-Tuberculosis			Malaria Prevention			HIV Prevention			Anti-Retroviral		
		K	G	LN	**K**	G	LN	**K**	G	LN	**K**	G	LN	**K**	G	LN
East	No	**0.067**	0.076	0.091	**0.067**	0.076	0.091	**0.058**	0.101	0.115	**0.045**	0.061	0.062	**0.045**	0.11	(31, 241)
	Yes	**0.059**	0.086	0.103	**0.056**	0.102	0.105	**0.056**	0.102	0.105	**0.55**	0.094	0.08			
South	Yes				**0.048**	0.097	0.113	†**0.039**	0.090	0.091				**0.066**	0.12	0.10
North	Yes	**0.033**	0.065	0.054	**0.029**	0.096	0.08	**0.029**	0.096	0.08	**0.059**	0.097	0.091			
	No	‡**0.052**	0.076	0.07	**0.041**	0.079	0.141	‡**0.052**	0.076	0.07				**0.043**	0.081	0.094
South	No	**0.041**	0.129	0.16				†**0.039**	0.090	(0.091	**0.050**	0.065	0.123	**0.049**	0.133	0.209
West & Central	No	**0.057**	0.104	0.209	**0.057**	0.104	0.209	**0.057**	0.104	0.209	**0.063**	0.095	0.074	**0.049**	0.231	0.093
	Yes	**0.042**	0.084	0.093	**0.043**	0.081	0.079	**0.043**	0.081	0.079	**0.036**	0.059	0.055	**0.047**	0.230	0.147

Notes: Fraction of 1000 random out-of-sample validations for which the null hypothesis that the out-of-sample data is distributed according to the in-sample estimated distribution can be rejected at 5% significance level using the Kolmogorov-Smirnov test. Results reported are for the kernel density estimation (K) and maximum likelihood estimation of in-sample gamma (G) and log normal (LN) distributions

average 20 more days than non-landlocked ones). The results also suggest that the unpredictable variability of PLT within categories is substantial (average coefficient of variation is 0.657).

Table A3 provides out-of-sample prediction accuracy of distributional forecasts for procurement lead time using three different methods of in-sample estimation: kernel density estimation (K); maximum likelihood estimation of in-sample gamma distribution (G); and maximum likelihood estimation of in-sample log normal distribution (LN). Since the kernel density estimation method outperformed the other two in each category, we chose it as the method of in-sample estimation of PLT.

A.2 Disbursement Inter-Arrival Times (DIT)

Excluding 24 negative values, 38 disbursements from potentially discontinued C-rated grants, 133 disbursements to multi-national principal recipients, and retaining 90- and 180-day reporting frequencies left 1658 DIT observations that we used for estimation and validation purposes (97% of the 2068 disbursements in the dataset). The other reporting frequencies (e.g. 150 days) were of small incidence and did not allow for a reliable analysis of disbursements within the same reporting frequency. 90-day grants account for 48% of disbursements in the resulting dataset, and 180-day grants account for 52%, respectively. Note that the number of DIT observations is smaller than the number of disbursements in the data since each DIT uses two disbursement data points.

We used multiple regression model specifications to identify drivers of DIT variability (see Table A4), and grouped disbursements into 40 categories characterized by region (Eastern, Northern, Southern and West & Central Africa), grant rating in previous reporting period (A, B1, B2, NR [non-rated] and Phase I) and period length (90 or 180 days). The methods for merging categories, constructing and assessing distributional forecast of DIT were identical to those used for PLT. Kernel density estimation resulted in a maximum of 63‰ repetitions rejected at the 5% significance level across all categories (mean 42‰). Prediction accuracy of the four categories with less than 5 data points was not assessed (baseline stockout levels in Table 4 changed by less than 0.02% when these categories were removed).

Multiple regression specifications (1)-(4) show the significance of Phase I, previous grant rating in Phase II, African region, a linear monthly time trend since the first disbursement of the grant and grant disease in predicting DIT. African region was a significant predictor of DIT in all specifications for all period lengths, as well as a subset of Phase I and the previous period grant rating in Phase II. The linear time trend was not a significant predictor in specification (2). The grant disease was also not a significant predictor, even before controlling for the previous grant rating in Phase II (specification 4). Therefore, specification (1) was selected for the predictive model of DIT on the basis of its associated number of significant predictors. The interpretation of the relatively low R^2 is the same as for PLT above.

Summary statistics of historical DIT per category in Table A5 exhibit substantial variations across prior grant rating R_k^i and geographic region $region(j)$. On average, DITs in East Africa

Table A4: Multiple regression model specifications for disbursement inter-arrival times

DIT	(1) 90-day	(1) 180-day	(2) 90-day	(2) 180-day	(3) 90-day	(3) 180-day	(4) 90-day	(4) 180-day
Phase I	-4.53	33.04	-5.42	33.38	-4.76	32.79	9.86	30.42
	(7.24)	(8.18)***	(6.17)	(7.73)***	(7.27)	(8.21)***	(7.11)	(6.78)***
Phase II Grant Rating (previous period) *Baseline: A*								
B1	5.22	-24.05	5.07	-23.81	4.34	-23.4		
	(8.17)	(7.56)***	-7.67	(7.57)***	-8.26	(7.60)***		
B2	-2.39	-16.2	-2.71	-15.38	-3.32	-14.57		
	(9.08)	(11.93)	(9.34)	(11.97)	(9.15)	(12.03)		
NR	80.51	-11.98	80.45	-11.51	80.11	-11.16		
	(13.50)***	(15.58)	(13.34)***	(15.56)	(13.47)***	(15.59)		
African region *Baseline: West & Central*								
North	33.88	27.46	32.28	24.92	33.49	25.71	23.47	31.57
	(8.77)***	(8.89)***	(8.82)***	(8.98)***	(8.96)***	(8.74)***	(8.64)***	(9.08)***
East	44.93	40.86	44.13	42.64	45.16	40.91	39.89	45.91
	(12.07)***	(8.03)***	(11.89)***	(7.91)***	(12.28)***	(8.13)***	(8.15)***	(12.23)***
South	16.42	22.95	15.16	18.27	13.89	22.28	20.01	15.46
	(7.85)**	(8.39)***	(8.46)**	(8.52)***	(8.06)*	(8.39)***	(8.22)**	(8.44)*
Linear Monthly Trend			-2.67	-1.27				
			(2.04)	(1.33)				
GF Disease *Baseline: Tuberculosis*								
HIV/AIDS					-2.13	-5.17	-6.73	-5.57
					(7.5)	(7.78)	(7.8)	(7.76)
Malaria					-1.95	-11.23	-12.4	-5.9
					(8.59)	(8.19)	(8.11)	(8.7)
Observations	814	844	814	844	814	844	814	844
R-squared	0.103	0.069	0.106	0.07	0.105	0.071	0.058	0.036

Notes: Multiple regressions of DIT for 90- and 180-day reporting grants on African region, grant Phase and: (1) previous period grant rating; (2) previous period grant rating and a linear monthly trend; (3) previous period grant rating and GF diease; (4) GF disease. Robust standard errors in parentheses; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A5: Summary statistics of disbursement inter arrival times (DIT) in days

Number of observations, mean and (5%, 95%) quantiles of DIT in each category; categories were merged whenever the Kolmogorov-Smirnov Test of common underlying distribution had a p-value larger than 0.5; merged categories are identified by either an omitted cell border or †.

		Grant Rating r^i_{k-1}														
	World Region region(j)	A			B1			B2			NR			Phase I		
		n	Mean	Quantiles	n	Mean	Quantiles	n	Mean	Quantiles	n	Mean	Quantiles	n	Mean	Quantiles
90-day Reporting Grants $l = 90$	West & Central	32	130	(35. 308)	†64	139	(28, 367)	19	146	(77, 222)	356	149	(39, 321)	356	149	(39, 321)
	North	17	184	(61, 359)	15	163	(42, 349)				2	275	(177, 373)	261	174	(44, 360)
	South				†64	139	(28. 367)				4	117	(76, 156)			
	East	3	261	(168, 335)					1	334				56	186	(28, 367)
180-day Reporting Grants $l = 180$	North	24	178	(34, 357)	64	179	(61, 354)	23	125	(20, 223)	23	125	(20, 223)	53	238	(67, 383)
	West & Central	57	159	(68, 294)	132	154	(33, 270)	11	135	(27, 211)	11	124	(40, 333)	95	192	(58, 369)
	South	77	231	(124, 377)				31	177	(26, 348)	23	154	(46, 323)	53	207	(66, 368)
	East				63	185	(35, 340)	11	260	(74, 391)	21	180	(31, 362)	103	206	(49, 371)

are 40–45 days longer than in West & Central Africa; DITs of 180-day grants in Phase I are 33 days longer than in Phase II; and DITs of 90-day non-rated grants in Phase II were 80 days longer than those rated A. Unpredictable variability of DIT within each category is also substantial (mean coefficient of variation is 0.508). Consistent with previous independent observations (Aidspan 2005), mean DIT is larger than grant reporting period in 19 out of the 24 categories, including all categories with 90-day reporting periods. Finally, no significant time trend of DIT during a grant lifecycle is observed.

Table A6: Distributional forecasts of disbursement inter-arrival times: out-of-sample prediction accuracy

	Grant Rating														
	A			B1			B2			NR			Phase I		
90-day Reporting Frequency															
African Region	**K**	G	LN	**K**	G	LN	**K**	G	LN	**K**	G	LN	**K**	G	LN
West & Central	**0.035**	0.048	0.048	†**0.043**	0.066	0.054				**0.044**	0.092	0.111	**0.044**	0.092	0.111
North	**0.039**	0.086	0.078	**0.042**	0.056	0.061	**0.043**	0.147	0.144				**0.057**	0.101	0.117
South				†**0.043**	0.066	0.054									
East													**0.044**	0.083	0.129
180-day Reporting Frequency															
North	**0.038**	0.145	0.087	**0.052**	0.81	0.87	**0.038**	0.129	0.118	**0.038**	0.129	0.118	**0.044**	0.069	0.061
West & Central	**0.046**	0.108	0.112	**0.044**	0.187	0.232	**0.027**	0.084	0.078	**0.063**	0.156	0.163	**0.050**	0.089	0.106
South	**0.043**	0.059	0.062				**0.053**	0.137	0.131	**0.035**	0.096	0.098	**0.054**	0.081	0.069
East				**0.050**	0.098	0.206	**0.041**	0.094	0.081	**0.031**	0.116	0.064	**0.046**	0.116	0.137

Notes: Fraction of 1000 random out-of-sample validations for which the null hypothesis that the out-of-sample data is distributed according to the in-sample estimated distribution can be rejected at 5% significance level using the Kolmogorov-Smirnov test. Results reported are for the kernel density estimation (K) and maximum likelihood estimation of in-sample gamma (G) and log normal (LN) distributions.

Table A6 provides out-of-sample prediction accuracy of distributional forecasts for disbursement inter-arrival times using three different methods of in-sample estimation: kernel density estimation (K); maximum likelihood estimation of in-sample gamma distribution (G); and maximum likelihood estimation of in-sample log normal distribution (LN). Since the kernel density estimation method outperformed the other two in each category, we chose it as the method of in-sample estimation of PLT.

A.3 Grant Rating Model Estimation

For each recipient i we focus on those grants, indexed by g, for which the recipient is i, i.e. $recipient(g) = i$.

We define the frequency of principal recipient i receiving a rating r by $frequency(i, r)$ and the transition probability from rating r_1 in period k to rating r_2 in period $k+1$ for any k by $probability(i, r_1, r_2)$ as follows:

$$
\begin{aligned}
frequency(i,r) &= \frac{\sum_{k=1}^{l} \sum_{\substack{g: \\ recipient(g)=i}} 1_{\{r_k^g = r\}}}{\sum_{\rho} \sum_{k=1}^{l} \sum_{\substack{g: \\ recipient(g)=i}} 1_{\{r_k^g = \rho\}}} \\
probability(i,r_1,r_2) &= \frac{\sum_{k=1}^{l-1} \sum_{\substack{g: \\ recipient(g)=i}} 1_{\{r_k^g = r_1\}} 1_{\{r_{k+1}^g = r_2\}}}{\sum_{\rho} \sum_{k=1}^{l-1} \sum_{\substack{g: \\ recipient(g)=i}} 1_{\{r_k^g = r_1\}} 1_{\{r_{k+1}^g = \rho\}}}
\end{aligned}
$$

Note that $probability(i, \text{Phase I}, r_2) = 0$ for all $r_2 \neq$ Phase I by definition of Phase I and Phase II.

B Model Assumptions

B.1 Qualitative Model Discussion

An important assumption of our model is that the amount of demand for medicines that may be satisfied by a given disbursement amount (*funding-to-demand ratio*) is constant and deterministic. That is, we assume that potential changes in demand and/or prices of health products potentially affecting the demand coverage associated with given funds are predictable and accounted for in the disbursement amounts. Although much attention and expertise is, in principle, dedicated to determining and reviewing the grant amounts requested from the Global Fund (see §1.1), in practice this assumption may or may not be perfectly satisfied. An important related observation, however, is that the demand for health products that is relevant to Global Fund grants occurs at the national

level and therefore exhibits limited variability because it is obtained by pooling demand across multiple geographic regions of an entire country. In addition, that demand tends to be unaffected by sudden local epidemics (e.g., cholera, yellow fever, polio), because of limitations in the categories of products that can be purchased with Global Fund grants as well as the time flexibility of these grants, so that other short-term funding mechanisms are typically used instead to fight these.

We assume unlimited shelf life for the product. While this is certainly a simplification of reality, it seems a good approximation when the funding-to-demand ratio f is relatively small and the central warehouse in the country has an efficient procedure of shipping inventory to the peripheral outlets. To ensure that our results are robust to the omission of shelf life, we implemented a model where each unit of a product would expire 18 months after procurement order placement, which we estimate to be the shortest product shelf-life in our dataset based on the information provided by manufacturers of these health products. The resulting stockout levels are within 0.1% of the ones with unlimited shelf-life for 90-day reporting grants and funding-to-demand ratio of $f < 2$, which seems reasonable in the current environment with relatively scarce funding. The same difference for 180-day reporting grants is 1.2%. Hence, our model underestimates stockout risks for 180-day reporting grants. However, given the large difference between stockout levels for 90- and 180-day reporting grants and the similar effect of product expiry on all 180-day reporting grants, this extended model would not change any of the existing paper conclusions.

Furthermore, demand for health products supported by some Global Fund grant component may in some cases span multiple medicines with possible interactions among them. In contrast, our model considers a single product at a time. Tender-related delays, which our model ignores, may also occur between fund disbursement and order placement. We believe that all assumptions just mentioned lead to underestimating stockout risks. DITs and PLTs are also assumed to be independent from the inventory level, which ignores the possibility that specific actions by Global Fund or principal recipients when inventory levels are low (e.g., higher priority of disbursement request, expedited transportation) could reduce them. An argument in support of this assumption is that the Global Fund does not currently have centralized visibility of the inventory levels of relevant health products in recipient countries, so that its ability to rationally prioritize in the short term between different disbursement or procurement requests may be limited. Because some reliable inventory level information may still be communicated to the Global Fund in an ad-hoc manner,

this assumption may still result in an over-estimation of stockout risks.

This model also implicitly assumes that countries do not have access to alternative sources of funding when gaps in Global Fund grant disbursements occur. In principle, countries may be able to access emergency funds through temporary reallocations between different budget lines or between grants from different donors. In practice however, it is suggested in Kraiselburd and Yadav 2013 that these emergency fund reallocations are limited because of constraints linked to fund traceability and transparency. In addition, in the next section we discuss some data suggesting that available sources of funding for health programs other than Global Fund grants are often limited in many low-income countries.

B.2 Single Funding Source

In order to provide some context for the limitation of our model scope to funding obtained from Global Fund grants, we discuss below the available data on donor funding sources for each of the three diseases. Figure B1 provides a graph of the proportion of donor funding from the Global Fund by country and disease. For each disease, the specific data provided is as follows:

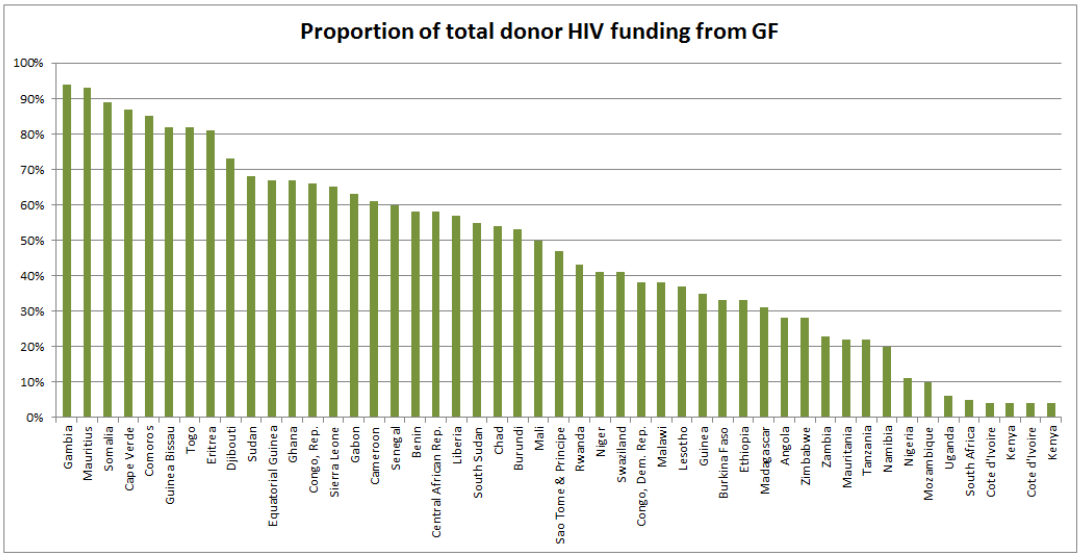
HIV/AIDS

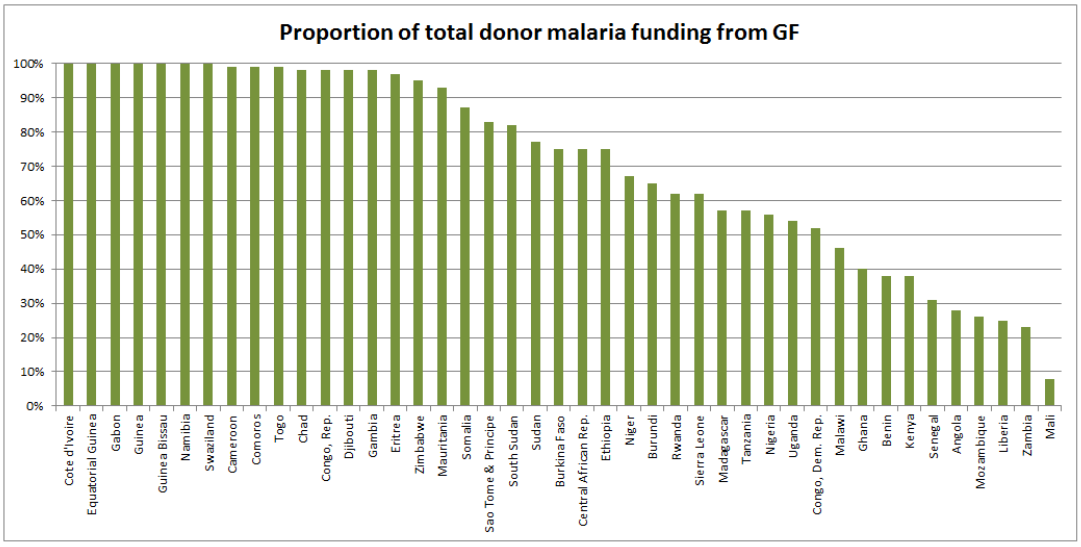
The Global Fund has been the single largest donor in 2009-2011 for 30 countries in Sub-Saharan Africa. Within those 30 countries, the Global Fund's support as a fraction of total donor funding ranges between 37% in Lesotho and 94% in Gambia.

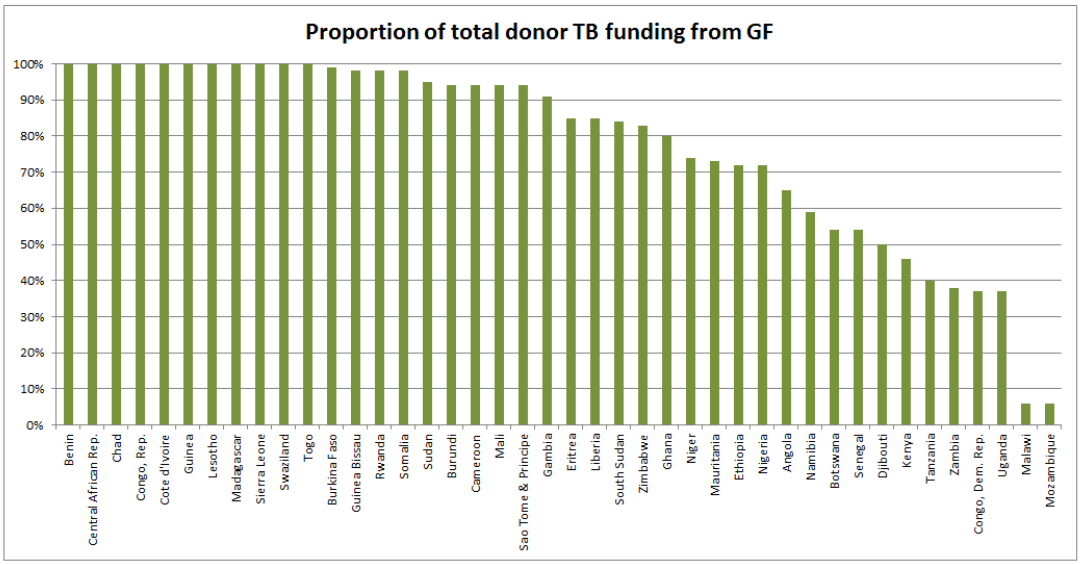
Malaria

We used data from the World Malaria Report (WHO 2011). The Global Fund accounts for approximately 50% of international financing of malaria control. In Sub-Saharan Africa the number is slightly higher at 57% (The Henry J. Kaiser Family Foundation 2012). For the three African countries in the World Malaria Report (Ghana, Madagascar and Nigeria), government malaria expenditure per person at risk is at least four times lower than by the GF. Furthermore, around 55% of GF funding is used for insecticide-treated nets and animalarial treatment against only 15% of government expenditure on health products. Hence, we conclude that about half of funding for malaria-related procurement in Africa is from the GF, meaning that the Global Fund is the predominant financier of procurement in a number of African countries.

Figure B1: Proportion of donor funding from Global Fund by disease







Notes: Data source – The Henry J. Kaiser Family Foundation (Kaiser 2012). See text for discussion on each disease data.

Tuberculosis

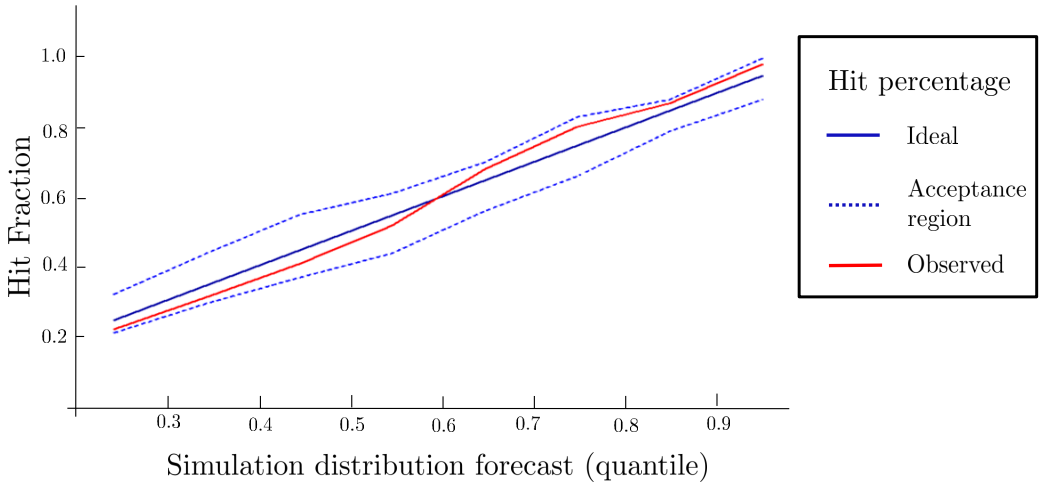
African governments account for 30% of their TB budgets and the GF is the primary donor financing TB programs(Figure B1). Hence, we conclude that the Global Fund is responsible for the procurement of most of TB-related drugs in Africa.

B.3 Model Validation Support

Despite the widely reported prevalence of national stockouts for essential medicines in Africa (see references cited in §1), we are not aware of any readily available dataset systematically documenting the historical national stockout episodes of any African country, let alone for all of Africa. Because of the public health impact of stockouts, this data is associated with management performance and political accountability, which makes it very sensitive and confidential.

Table 2 shows average simulated in-sample implementation lengths obtained with baseline parameters, against average actual in-sample and out-of-sample implementation lengths, and Figure B2 plots the hit fraction of out-of-sample observations of grant implementation lengths relative to the in-sample simulation distribution; both are discussed in §3.3.

Figure B2: Hit fractions of out-of-sample observations of grant implementation lengths relative to in-sample simulated distribution.



Notes: Hit percentages - fraction of out-of-sample observations - below θ quantile of the corresponding predicted implementation length distribution for θ between 0.25 and 0.95 in increments of 0.1; 90% acceptance region for the null hypothesis that out-of-sample observations were realizations from the corresponding distributional forecasts of the simulation.

C Additional Model Results

C.1 Causal Impact of Grant Monitoring Frequency

The results in §4.1 show that the estimated stockouts of *all* 90-day grants are substantially higher than the estimated stockouts of *all* 180-day grants. If these stockout risks were driven by underlying recipient risk factors as opposed to the reporting mechanism, this would imply that the Global Fund would have correctly assessed these risk factors in advance when determining the appropriate grant monitoring frequency for all 641 grants of our data set, a seemingly implausible scenario. In addition, there are 26 grants in our dataset for which the Global Fund decided to change the reporting frequency at some point during their implementation (15 grants switched from 90-day to 180-day, 11 grants from 180-day to 90-day). Such changes were likely deliberate and carefully thought out, as the remaining 435 grants kept the same reporting frequency throughout. These 26 grants thus provide an opportunity to consider modified paired counterfactual grants providing an indication of what would have happened if different performance monitoring mechanisms had been used, holding the underlying implementation risks constant. To that end, we simulated the corresponding 26 associated virtual grants with the exact same ratings history over time as their paired originals, but with symmetrically opposed reporting frequencies in each time period: each single reporting period of 180 days was replaced by two consecutive reporting periods of 90 days and each pair of consecutive 90 day periods was replaced with a single reporting period of 180 days. While our conclusions are limited by sample size, we observe that the mean difference between the simulated implementation lengths of the counterfactual and original grants is statistically higher (with 91.9% significance level) than the mean difference between the simulated lengths of the original grants and the corresponding actual implementation lengths. Hence, at 90% significance we cannot reject the hypothesis that the grant implementation lengths are driven by the performance monitoring mechanism itself rather than any underlying risk factors associated with each recipient.

C.2 Effect of Health Product Type on Stockouts

The results shown in Table C1 suggest that stockout differences among countries in the same African regions and between landlocked and non-landlocked countries are limited. While grant rating and countries' geographic type are important predictors of DIT and PLT, respectively, they do not affect stockout risks significantly. In North Africa and West & Central Africa, expected lost demand did

Table C.2: Simulated average proportion of demand lost over three years

	Health Product Type m									
	Anti-Malarial		Anti-Retroviral		Anti-Tuberculosis		HIV Prevention		Malaria Prevention	
African region &	Reporting Frequency (days) l									
Country j	90	180	90	180	90	180	90	180	90	180
North Africa										
Landlocked										
Chad	28.05	3.11	27.91	2.81	31.43	5.15	29.61	2.17	31.24	5.28
Mali	29.42	3.67	29.06	3.36	32.12	5.58	29.95	2.43	33.06	5.33
Niger	31.16	3.88	31.76	3.19	33.42	5.46	31.14	2.02	34.57	5.25
South Sudan	30.82	3.46	30.25	3.04	33.19	5.36	30.02	2.38	35.18	5.37
Non-landlocked										
Algeria	26.11	2.63	26.76	2.13	26.81	3.08	25.41	2.36	26.66	2.86
Djibouti	29.27	2.75	28.12	1.82	29.9	3.2	28.54	1.42	29.92	2.47
Egypt	29.92	3.05	30.33	1.94	32.17	3.81	30.18	1.83	32.12	3.14
Jordan			33.29	3.44	32.06	4.67	32.41	2.98		
Mauritania	31.37	4.64	32.28	3.97	33.62	4.81	32.47	2.7	34.61	4.56
Morocco			31.47	3.85	32.05	3.62	30.12	2.51		
Somalia	30.61	3.42	30.38	2.51	32.14	3.29	30.37	2.11	31.62	3.72
Tunisia			29.19	3.28	33.55	4.23	29.81	3.02		
Yemen	29.09	2.35	32.13	3.72	30.41	2.71	32.06	2.96	30.72	2.91
East Africa										
Landlocked										
Burundi	43.49	12.61	49.83	12.88	56.34	12.9	51.27	9.21	50.17	11.62
Ethiopia	42.61	9.04	40.14	9.25	48.13	11.17	40.15	8.93	49.31	12.17
Rwanda	39.97	10.18	38.65	10.09	44.27	10.05	45.83	8.14	44.53	11.48
Uganda	41.17	9.68	41.06	10.53	48.16	10.69	40.18	7.65	39.19	12.29
Non-landlocked										
Comoros	43.21	10.58	43.69	11.68			44.66	9.82	42.23	9.04
D.R. Congo	45.78	12.31	45.2	12.17	47.31	12.68	44.18	10.38	46.07	11.37
Eritrea	43.96	11.45	43.68	10.82	45.52	12.41	43.95	11.03	44.19	10.13
Kenya	46.57	12.74	53.21	11.94	54.26	12.9	53.08	11.64	54.28	10.97
Madagascar	46.79	10.18	46.37	12.18	47.2	11.65	46.79	11.92	52.31	7.14
Mauritius			54.37	9.02			52.8	8.29		
Tanzania	43.19	10.03	48.21	8.26	48.26	10.23	47.16	7.51	45.92	7.83
Zanzibar	40.22	11.66	42.05	10.31			41.83	10.03	39.7	8.27
South Africa										
Landlocked										
Botswana			33.26	4.13	41.57	5.78	32.72	3.91		
Lesotho			33.41	5.58	41.49	6.13	32.46	4.12		
Malawi	31.54	4.52	27.04	4.11			25.72	3.82	22.38	3.89
Swaziland	35.12	5.68	32.62	4.18	41.62	6.42	32.19	3.7	34.55	3.55
Zambia	32.39	4.12	28.71	3.78	40.35	6.68	27.12	3.41	32.2	4.25
Zimbabwe	32.7	6.45	30.47	4.82	41.7	6.35	29.77	4.14	30.61	4.61
South Africa										
Non-landlocked										
Angola	35.37	6.72	31.22	5.27	34.63	6.8	31.5	4.37	34.24	4.74
Mozambique	27.65	5.79	25.38	4.24	34.08	5.56	22.02	3.87	24.64	4.28
Namibia	29.37	5.37	26.34	4.45	27.62	5.41	25.17	3.65	27.57	4.8
South Africa			28.26	4.71			27.16	3.3		
West &Central										
Landlocked										
Burkina Faso	22.61	4.13	24.05	4.38	21.31	2.29	22.32	1.39	21.38	2.9
CAR	26.74	2.63	25.81	3.15	26.74	2.37	24.27	1.11	27.16	2.27
Non-landlocked										
Benin	25.29	4.52	23.25	1.69	22.38	2.19	22.47	1.26	25.18	4.38
Cameroon	23.1	2.8	22.06	3.14	24.4	3.45	22.52	2.65	24.62	3.55
Cape Verde	21.38	2.16	17.83	2.19			18.06	1.93	22.45	2.18
Congo	19.26	2.76	22.45	1.23	18.83	1.79	23.26	1.36	18.32	2.63
Cte d'Ivoire	24.72	2.04	24.58	1.69	22.68	1.63	23.61	1.62	25.59	2.03
Guinea	22.05	2.17	21.29	1.75	22.15	2.31	21.46	1.41	22.36	2.2
Equat. Guinea	24.17	2.08	21.57	2.02			22.73	1.75	23.71	2.16
Gabon	18.29	2.74	18.04	2.5			18.16	1.8	18.48	2.72
Gambia	20.16	2.28	18.58	1.31	22.84	2.17	17.03	1.38	21.14	2.36
Ghana	22.31	2.01	21.62	1.26	22.58	1.42	20.41	1.02	23.58	1.52
Guinea-Bissau	19.73	3.07	17.32	1.88	20.05	2.24	17.37	1.95	18.26	1.99
Liberia	18.57	2.89	17.14	2.13	19.26	2.87	18.13	1.89	19.48	2.73
Nigeria	21.42	2.52	22.45	1.92	23.78	2.18	21.69	1.74	21.53	2.19
Senegal	18.25	2.71	20.18	1.81	18.66	2.91	18.53	1.68	18.73	2.65
Sierra Leone	19.34	2.39	20.92	1.74	21.47	2.53	20.03	1.72	19.96	2.78
Togo	18.47	2.4	19.11	2.03	19.75	2.78	18.27	1.79	22.62	2.69

Notes: Simulated average proportion of demand lost over three years by world region and country, type of drug and reporting frequency under baseline assumptions (percentage). Results based on 5,000 simulated replications yielding a 95% CI length less than 0.8% of the estimated mean estimate in each scenario. Results from several PRs in same country were weighted by relative volume of approved GF grant amounts. Scenarios with no grant approved historically for a given country and disease were not simulated.

not vary significantly across health product types. However, anti-TB drugs experience significantly higher stockouts only in landlocked countries in the regions of East and South Africa. This is the result of longer PLTs for those countries, which combined with DITs lead to higher stockouts. Hence, the question here is why procurement lead time for anti-TB drugs is higher in those countries?

This is not specifically due to the geographic region or landlock situation, as this effect is not observed for landlocked countries in North and West Africa. Rather, conversations with current and former Global Fund staff point out to a specific supplier practice consisting of pooling orders from these countries. This could indeed lead to longer PLTs since the supplier might wait for more orders before processing and distributing the one(s) already received. This explanation is also supported by some additional analysis of the procurement data showing that the 10 countries with substantial higher stockouts for anti-TB products are 38% more likely to share drug suppliers than any other countries. These 10 countries are also 72% more likely to share suppliers among themselves than with another country.

C.3 Yearly Replenishment

As a robustness check, we also performed the simulation with yearly procurement instead of every period, that is procurement every two periods for 180-day reporting grants and every four periods for 90-day reporting grants (Table C2).

Table C2: Expected stockouts by African region under yearly replenishment (percentages)

		Yearly Replenishment Policy									
		Phase I		Phase II		IR		BF		SF	
	Reporting Frequency (days)	90	180	90	180	90	180	90	180	90	180
African Region	North	22.6%	1.0%	20.8%	3.1%	16.2%	0.7%	19.5%	3.0%	9.9%	1.06%
	East	46.1%	5.2%	35.2%	8.4%	29.3 %	2.9%	35.0%	7.7%	22.7%	2.78%
	South	22.2%	2.1%	20.3%	3.0%	14.8%	1.0%	18.2%	2.8%	7.1%	0.86%
	West & Central	18.5%	1.7%	17.1%	2.2%	12.0%	0.4%	15.6%	2.0%	4.4%	1.15%

Notes: Expected stockouts for the proposed base case interventions by African region over a three-year time horizon: 180-day sufficiency of initial inventory and cash buffer level of 10%; results based on 5,000 replications ensuring the length of the 95% confidence interval is less than 1% of the estimated expected stockout in each scenario.

The interpretation of the per period cash buffer level is unchanged, so that the yearly replenishment buffer was twice as large as the period one buffer for 180-day reporting grants and four times as large for 90-day reporting grants. For 90-day reporting grants, stockouts with yearly replenishment

were smaller compared with the corresponding per period replenishment: 5–10pp for the Phase I and Phase II scenarios; 4–6pp for IR; 1–3pp for BF and 2–5pp for SF. Yearly replenishment did not lead to a statistically significant reduction of expected stockouts for 180-day reporting grants.

C.4 Probability of Stockout Over Time

Table C3 provides the implied probability of a stockout (percentage) over three years. Strikingly, programs with 90-day reporting grants had on average at least 50% chance of a stockout in all regions, with the probability raising to 98.6% in East Africa.

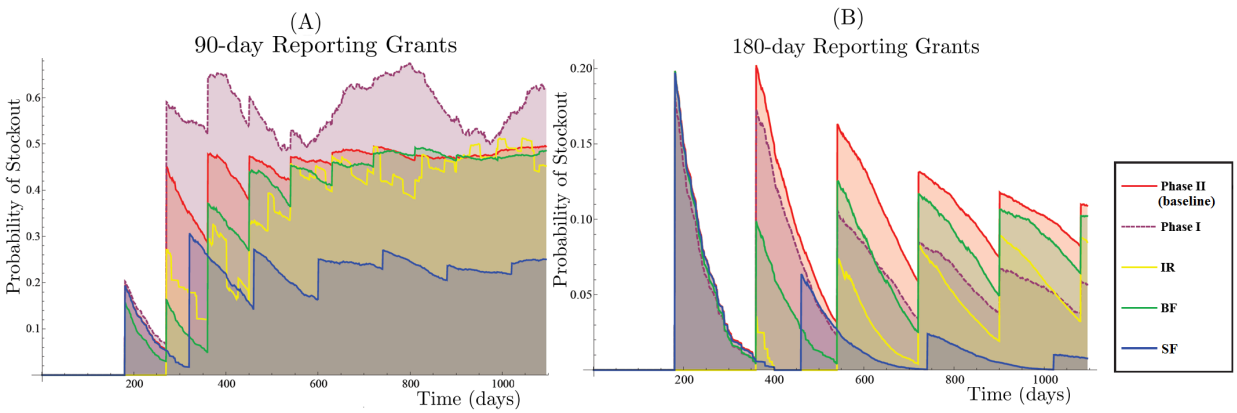
Table C3: Probability of a stockout occurrence over three years (percentage)

		Policy									
		Phase I		Phase II		IR		BF		SF	
	Reporting Frequency (days)	90	180	90	180	90	180	90	180	90	180
African Region	North	68.8%	10.2%	68.7%	19.7%	56.7%	6.9%	64.2%	14.4%	42.1%	9.7%
	East	98.6%	26.1%	76.6%	35.3%	71.5%	19.3%	74.1%	31.9%	65.3%	19.1%
	South	68.5%	14.9%	59.7%	19.5%	52.6%	9.9%	56.6%	13.9%	37.4%	7.6%
	West & Central	57.2%	14.3%	50.4%	15.1%	41.8%	6.1%	45.3%	12.1%	31.5 %	7.5%

Notes: Percentage of 5,000 replications in which stockout over three years was positive; results across different drugs and within African region weighted by the number of reported malaria cases in 2010 for anti-malaria and malaria prevention drugs; sum of people in need of and on ARV treatment for ARV drugs; number of people living with AIDS for HIV prevention drugs; new TB cases in 2011 for anti-TB drugs (The Henry J. Kaiser Family Foundation. 2012).

Plotting the probability of stockout on a particular day in Figure C1, the probability of stockout achieves a local minimum at times which are multiples of the period length following the exhaustion of initial inventory. The intuition for 90-day reporting grant is as follows: since initial inventory suffices for 180 days, no stockout before that time occurs; stockout between day 180 and 270 occurs if any only if no procurement order has been received; this probability decreases over time and thus so does the probability of stockout; the sharp increase on day 271 stems from the fact that stockout no longer occurs if and only if no procurement order has been received, but it could also occur if a procurement order has been received before day 180 so that it was used to satisfy demand between day 180 and 270; as long as no second order has been received, stockout at day 271 occurs. The interpretation of these results for 180-day grants is similar.

Figure C1: Probability of stockout for 53 African countries over time

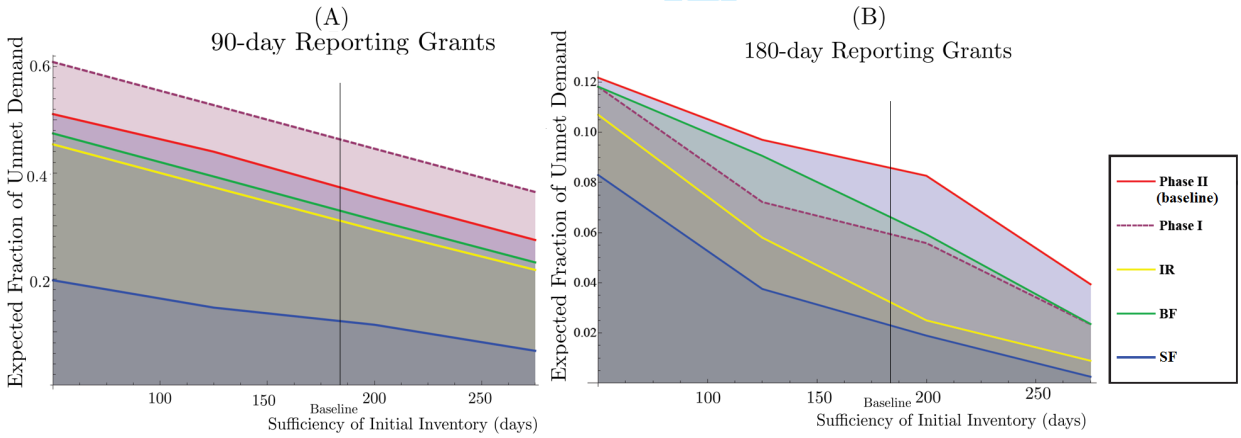


Notes: Simulated probability of stockout aggregated over 53 African countries over a three-year time horizon. Probability of stockout in Africa, weighted sum of fraction of 5,000 scenarios in which inventory was zero; results for African region weighted by estimated proportional demand; initial inventory sufficient for 6 months; cash buffer level of 10% for all scenarios except synchronized financing where it is 50% by definition ; scenarios are Phase II (baseline), Phase I, IR - instantaneous replenishment, BF - bridge financing and SF - synchronized financing.

C.5 Initial Inventory

Figure C2 shows the effect of initial inventory on the expected fraction of lost demand over three years for all policy scenarios.

Figure C2: Effect of initial inventory on lost demand



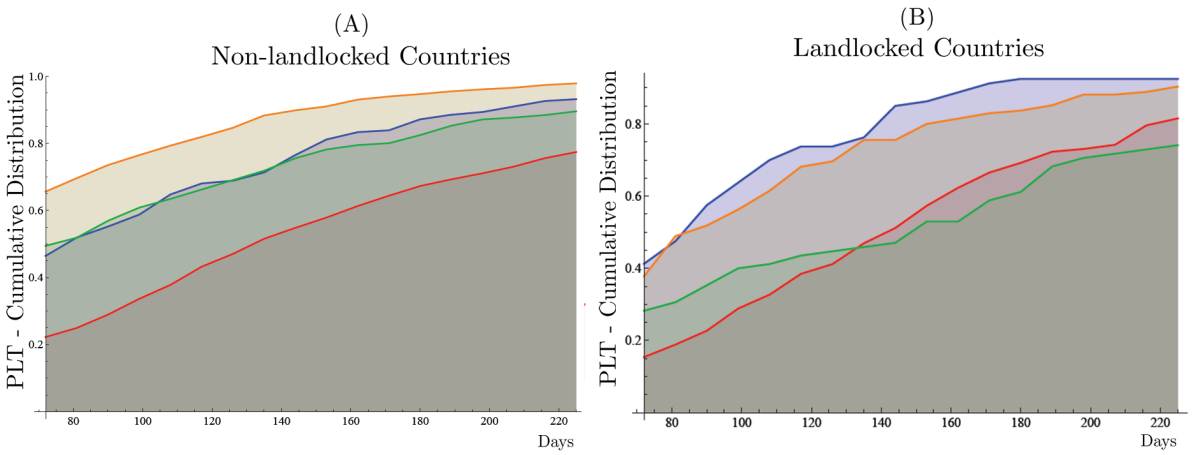
Notes: Effect of initial inventory on simulated average proportion of lost demand aggregated over 53 African countries over a three-year time horizon. Aggregate expected stockout in fraction of lost demand in Africa with results for each African region weighted by estimated proportional demand; time horizon is three years; 0% cash buffer level for all scenarios except synchronized financing where it is 50% by definition; scenarios are Phase II (baseline), Phase I, IR - instantaneous replenishment, BF - bridge financing and SF - synchronized financing.

D Additional Data Analysis

D.1 Cumulative Probability Distributions

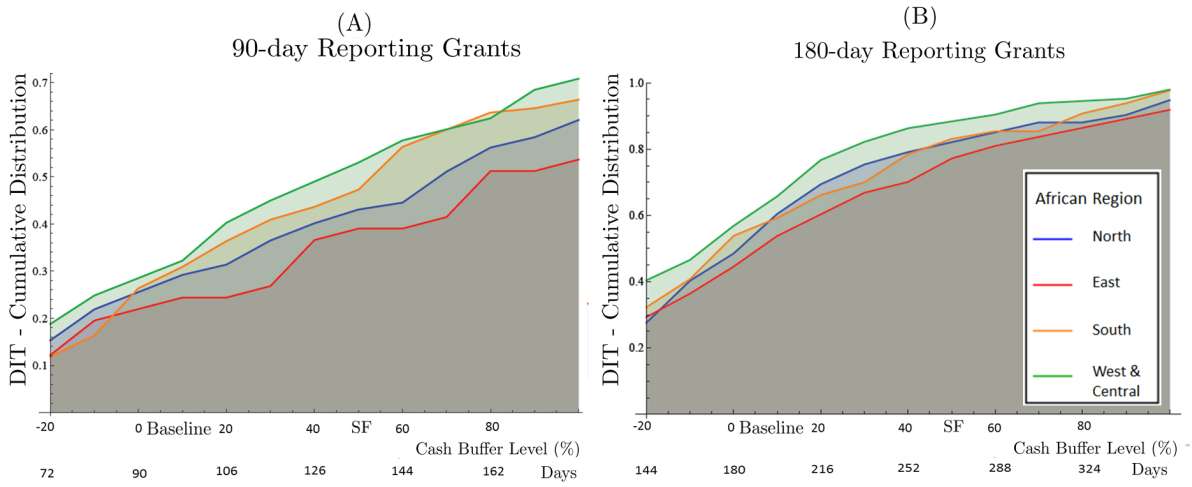
Figures C3 and C4 provide the cumulative distributions of procurement lead time and disbursement inter-arrival time, respectively, by African region. The lower cumulative distributions in East Africa in all instances exemplify the qualitative results discussed in §4.2.

Figure C3: Cumulative probability distribution of procurement lead time distribution



Notes: Estimated cumulative probability distribution of procurement lead-time distribution. Fraction of PLT observations in each African region not exceeding a fixed number of days.

Figure C4: Cumulative probability distribution of disbursement inter-arrival times



Notes: Estimated cumulative probability distribution of disbursement inter-arrival times in relation to cash buffer level. Fraction of DIT observations in each African region not exceeding the sufficiency of the per period disbursement, defined as period length plus cash buffer level.

D.2 Grant Reporting Frequency

We estimated some additional regression models to further investigate the drivers of grant reporting frequency. All specifications indicate that the grant phase is the most useful predictor of reporting frequency: grants in Phase II are around 30% more likely to have 180-day reporting frequency than grants in Phase I. We also included three variables related to the trust between the Global Fund and the PR: the number of past PR grant ratings of a certain kind, the number of years since the Global Fund first disbursed to the PR and the number of approved PR grants. All of these are historical data up to the reporting frequency observation, e.g. for a reporting frequency observed in March 2012 we count the relevant past disbursements and grants up to March 2012.

Each grant approved by the Global Fund increases the probability of 180-day reporting frequency of future grants by around 8% (statistically significant at the 95% level). There is also a positive correlation between past PR performance and reporting frequency: specification (3) shows a larger number of past A ratings increases the probability of 180-day reporting frequency, while past C ratings decreases it. Looking at the time of collaboration alone, the number of years since first disbursement, we do not observe a significant effect on reporting frequency.

Our analysis shows that there is a correlation between trust and longer reporting frequency and this trust is based on past performance (ratings, grant approvals) rather than duration of the collaboration. In specification (5), we also control for the grant disbursement amount and the PR country corruption perception index.

The Global Fund classifies its principal recipients into Government (Ministry of Health or others), Multinational Organizations (e.g. UNDP, the Red Cross) and NGOs (local organizations, such as Mission East). Regression specifications (2)-(5) show strong correlation between PR type and grant reporting frequency: multinational organizations tend to have longer reporting frequencies than government and NGOs tend to have shorter.

Reporting frequency *Baseline: 90-day*	(1)	(2)	(3)	(4)	(5)
Grant Phase *Baseline: Phase I* Phase II	0.31 (0.07)***	0.34 (0.09)***	0.27 (0.08)***	0.26 (0.08)***	0.28 (0.07)***
PR Type *Baseline: Government* Multinational Organization		0.19 (0.04)***	0.11 (0.03)***	0.09 (0.04)**	0.07 (0.03)**
NGO		-0.09 (0.04)**	-0.10 (0.04)***	-0.11 (0.04)***	-0.11 (0.03)***
Number of past PR ratings: A			0.05 (0.014)*		
B1			0.01 (0.02)		
B2			-0.04 (0.03)		
C			-0.14 (0.04)***		
Years since first disbursement to PR			0.07 (0.05)		0.08 (0.05)
Number of approved PR grants			0.08 (0.03)**		0.09 (0.04)**
Disbursement Amount (Million USD)				-0.03 (0.01)***	-0.03 (0.01)***
PR Country Corruption Index				0.02 (0.007)***	0.03 (0.005)***
Observations	1,658	1,658	1,658	1,658	1,658
R-squared	0.083	0.091	0.105	0.098	0.094

Notes: Multiple regressions of reporting frequency on grant phase, PR type, past PR ratings, years since first disbursement to PR, number of approved PR grants, grant disbursement amount (million USD) and PR country corruption perception index by Transparency International. Robust standard errors in parentheses; $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

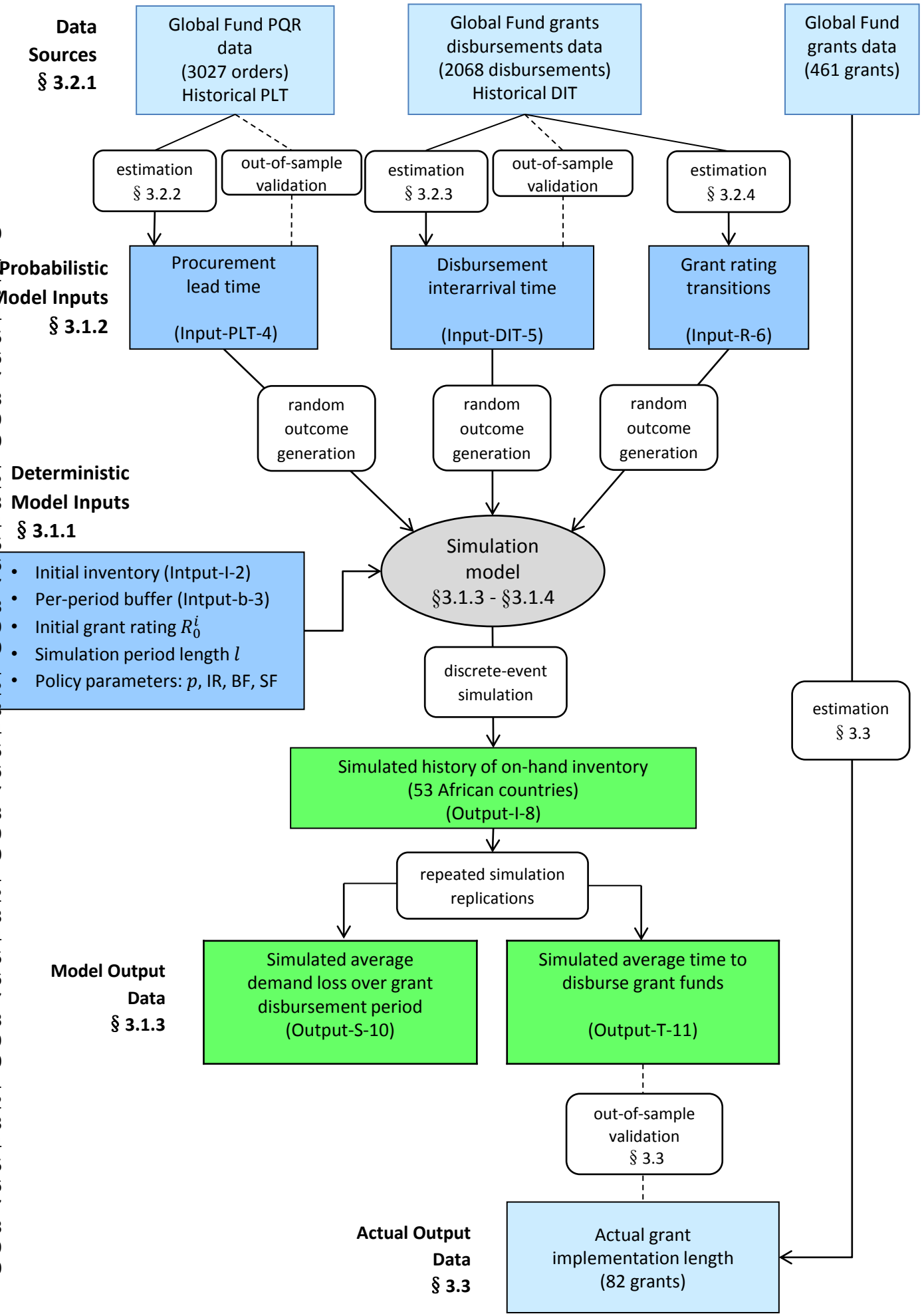
D.3 Corruption

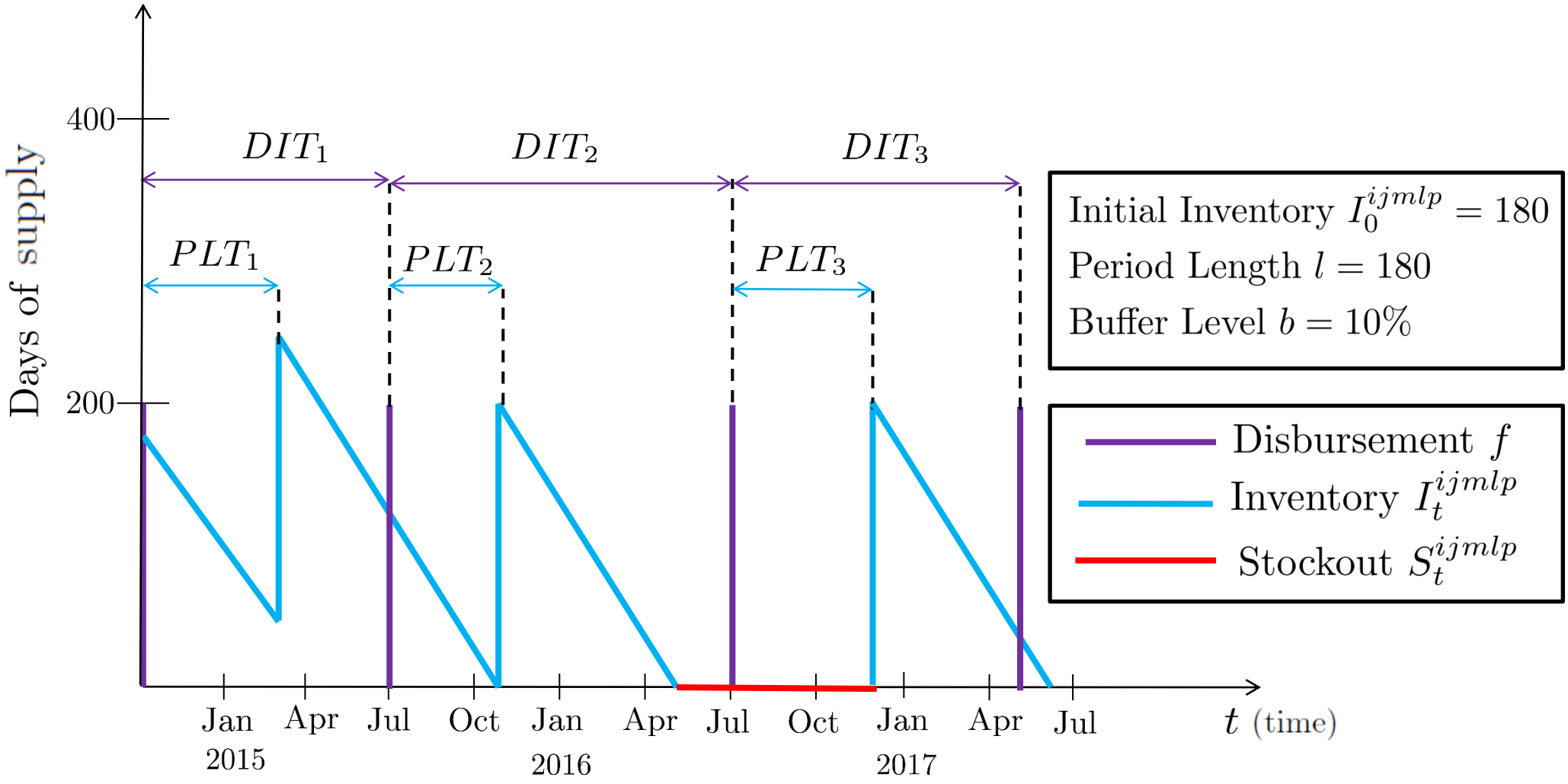
Higher corruption perception index (higher values suggest less corruption) is associated with shorter DITs for 90-day grants, but the effect is not significant for 180-day reporting grant. This observation

can be used by the Global Fund to better understand the association between country characteristics and expected stockouts.

DIT	90-day	180-day
Controls Phase Indicator Grant Rating African region Linear Monthly Trend GF Disease		
Corruption Perception Index (0 to 100)	-6.73 (1.82)***	-4.35 (2.43)
Observations	814	844
R-squared	0.105	0.070

Notes: Multiple regressions of DIT on corruption perception index by Transparency International controlling for all variables from Table A3. Robust standard errors in parentheses; ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.


Data Sources § 3.2.1
Global Fund PQR data (3027 orders) Historical PLT
Global Fund grants disbursements data (2068 disbursements) Historical DIT
Global Fund grants data (461 grants)
estimation § 3.2.2
out-of-sample validation
estimation § 3.2.3
out-of-sample validation
estimation § 3.2.4
Probabilistic Model Inputs § 3.1.2
Procurement lead time (Input-PLT-4)
Disbursement interarrival time (Input-DIT-5)
Grant rating transitions (Input-R-6)
random outcome generation
random outcome generation
random outcome generation
Deterministic Model Inputs § 3.1.1
• Initial inventory (Intput-I-2)
• Per-period buffer (Intput-b-3)
• Initial grant rating R_0^i
• Simulation period length l
• Policy parameters: p, IR, BF, SF
Simulation model §3.1.3 - §3.1.4
discrete-event simulation
estimation § 3.3
Simulated history of on-hand inventory (53 African countries) (Output-I-8)
repeated simulation replications
Model Output Data § 3.1.3
Simulated average demand loss over grant disbursement period (Output-S-10)
Simulated average time to disburse grant funds (Output-T-11)
out-of-sample validation § 3.3
Actual Output Data § 3.3
Actual grant implementation length (82 grants)


Days of supply
400
200
DIT_1
DIT_2
DIT_3
PLT_1
PLT_2
PLT_3
Jan
Apr
Jul
Oct
Jan
Apr
Jul
Oct
Jan
Apr
Jul
2015
2016
2017
t (time)
Initial Inventory $I_0^{ijmlp} = 180$
Period Length $l = 180$
Buffer Level $b = 10\%$
Disbursement f
Inventory I_t^{ijmlp}
Stockout S_t^{ijmlp}

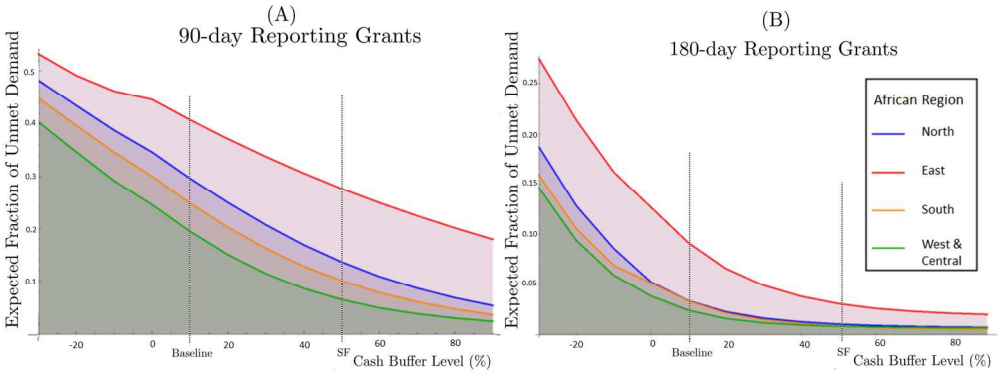


Figure 3: Fraction of lost demand for di
fferent cash bu
ffer levels

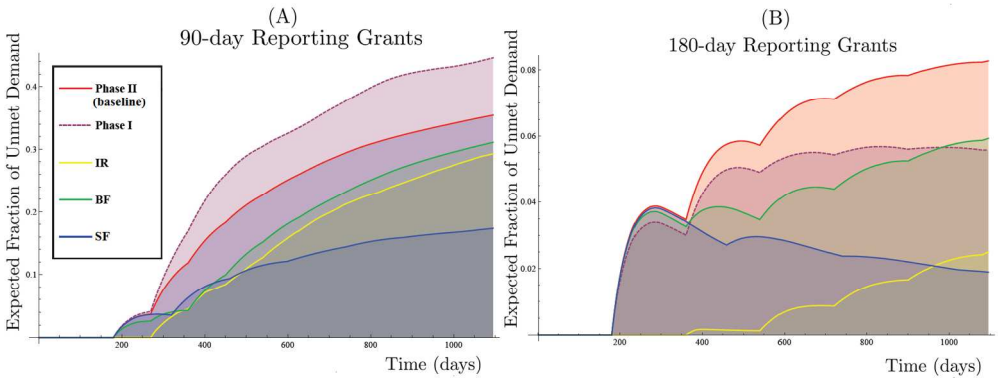


Figure 4: Fraction of lost demand over various time horizons

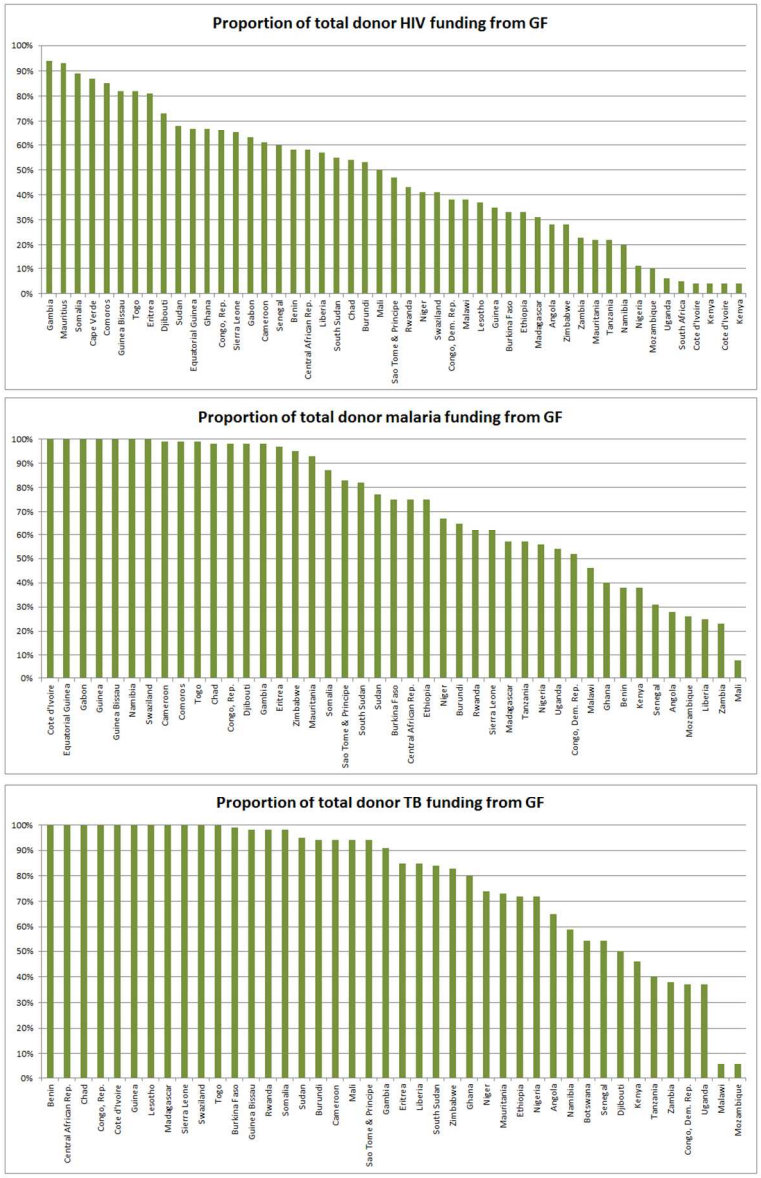
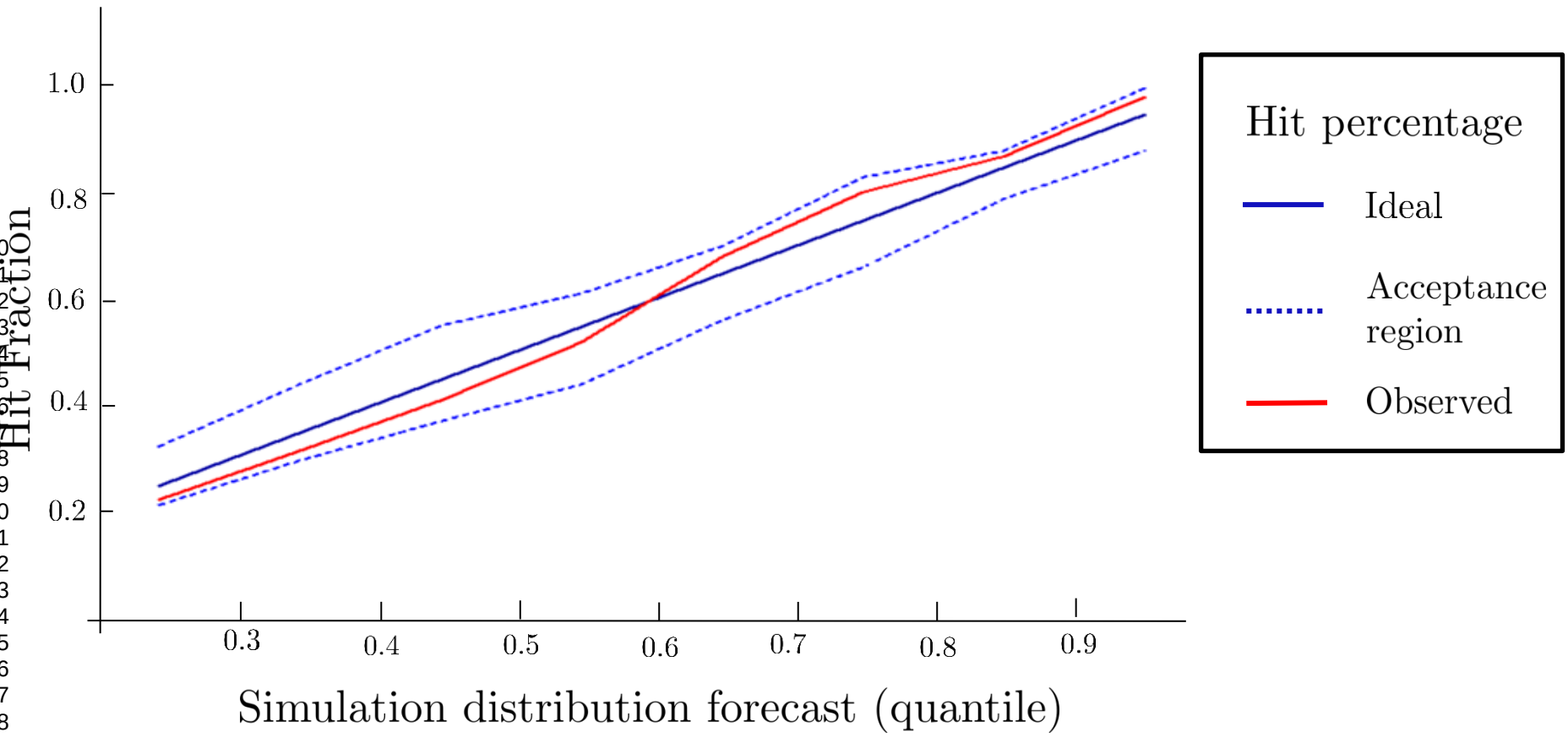


Figure B1: Proportion of donor funding from Global Fund by disease


Hit Fraction
1.0
0.8
0.6
0.4
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Simulation distribution forecast (quantile)
Hit percentage
Ideal
Acceptance region
Observed

Editor in Chief

Dear Authors:

Greetings! We are using this impersonal format because we are copying this email to the referees.

Manuscript ID POM-Apr-15-OA-0267.R1 entitled "National Drug Stockout Risks and the Global Fund Disbursement Process for Procurement" that you submitted to Production and Operations Management, has been reviewed. The comments of the review team are included at the bottom of this letter.

First of all, I would like to thank you for considering Production and Operations Management to publish your work. I have read the paper and the comments of the review team. The review team has recommended some minor revisions to your manuscript. Based on my own reading of the paper and the comments of review team, I invite you to revise your manuscript.

[…]

Please submit the revision before the end of this month. Otherwise we may have to consider your paper for regular issues of POM Journal.

Once again, thank you for submitting your manuscript to Production and Operations Management and I look forward to receiving your revision.

Sincerely,
Prof. J. Shanthikumar
Production and Operations Management
shanthikumar@purdue.edu

Dear George,

Thank you very much for the time and efforts you spent overseeing the review of our manuscript.

In the following, we provide a detailed description of how we addressed the comments of all reviewers.

We hope you like the revised version of the paper.

Yours sincerely,
The authors

Senior Editor

Dear Author(s):

I have now received three reviews of your paper. All reviewers state they appreciate the efforts you made to revise the paper in line with their previous comments. I concur.

The verdicts are different though. While reviewer 1 recommends minor revision, reviewer 2 would like to see a major revision. Reviewer 3 recommends acceptance. My personal view is that the paper should be accepted after a minor revision. If you clearly indicate what changes were made in response to the comments of the reviewers, i shall not send the paper out again, but instead deal with it directly.

Reviewer 1 lists 3 major comments. I would like you to pay particular attention to the second one. The reviewer also lists a long set of minor comments which are easy to deal with. Reviewer 2, although suggesting major revision, really seems to list comments that i would call relatively minor. Nevertheless, please consider them carefully and try to respond.

I would also agree with reviewer 1's request to shorten the paper, especially the 'future research' part. However, any shortening of the paper where possible would benefit you in terms of clarity of contribution and efficiency.

Congratulations on. A great piece of work. Apologies for a rather long turnaround time.

Sincerely.

Thank you very much for the time and effort spend on the review of our manuscript, for your positive comments about our work and for summarizing the reviewers' comments.

As part of this revision we have carefully considered your comments and that of the referees, and have made accordingly a number of changes to the paper which are all described in details in the remainder of this response document. In particular, we have substantially reduced the length of the discussion section and our discussion of future research opportunities, as requested by Reviewer 1 and yourself.

We hope you like the revised version of our paper, and look forward to your feedback.

Reviewer 1

The authors have undertaken a very significant and good faith revision of the document based the feedback from the earlier submission. The exposition and problem statements are now presented in a much better manner.

Thank you very much for the time and efforts spent reviewing our revised paper, and for your positive comments about our work.

MAJOR:

* The abstract, introduction, and contribution all talk about factors which are important, but it is not systematic in saying in which way is it important or in which direction the important decision variables influence the outputs. For example, the paper enumerates 3 major questions and summarizes responses in the bottom half of page 3 and top paragraph of page 4. The third of these three points is clear that front loading of disbursement reduces expected stockouts. The first two of these three points only identify factors and do not provide insights as to the direction of the benefit. This type of observation occurs also in the abstract and the conclusions.

Thank you for pointing out this opportunity to clarify our results. In the statement of the first two research questions and related results summary as part of the introduction, we now explicitly mention the direction in which key variables influence the outputs, namely the link between higher stockout risks and higher reporting frequency and the higher stockout risks in Each Africa. The edited statement of these two research questions in the paper revision is as follows:

"1. What is the impact of the grant recipient performance monitoring frequency (i.e., the scheduled frequency of grant disbursements and reporting and monitoring activities) used by the Global Fund on the stockouts experienced by receiving countries? The results discussed in §4.1 suggest that grants with higher reporting frequency exhibit substantially higher stockout risks. Hence, over the first 11 years of the Global Fund, there was a clear effective trade-off between the extent of its performance monitoring activities and the effectiveness of the procurement funds it disbursed.

2. Are there some geographic patterns affecting the risks of stockouts experienced by Global Fund grant recipients? We find that the African regions used internally by the Global Fund for organizational purposes constitute a substantial driver of stockout risks variability, with grant recipients in East Africa facing significantly higher stockout risks than in other regions. This suggests that a substantial fraction of the stockouts facing grant recipients are driven by organizational features and specific processes used by the Global Fund, as opposed to underlying risk factors associated with these recipients (see §4.2)."

We have also added similar clarifications of these points in the abstract and the conclusion.

* The discussion section is too long. I'm not a big fan of long 'future work' sections. If the authors could discuss the main limitations of the model, and remove a lot of the 'future work', the paper would be more concise and hopefully more effective.

Following this suggestion and comments by the Senior Editor, we have removed section §5.2, which is now summarized in just one paragraph at the end of the paper.

* The model is much clearer thank you. In the results and discussions, the text is organized to elaborate more

fully about the three main questions. I like the overall structure. However, some of the discussion involves interpreting the empirical data and some around the model. The discussion was not always clear to me, and a tightening and focus would strengthen the paper. I'm hesitant to offer specific suggestions as I'm not sure of the best way to do this, hence I provide this general reaction.

Following this comment, we have moved the admittedly long discussion on causal analysis between reporting frequency and stockout risks to the appendix (§C.1 – p.12-13). In addition, we have also clarified the internal structure of each subsection §4.1-§4.3 at the beginning of section §4, specifically mentioning that within each of these subsections we first present relevant empirical results and then discuss their implications.

MINOR POINTS:

* page 8, line 3-4: is 'entirely' the right word? 'entirely' seems very strong.

We think that the use of this word in that sentence is accurate, because it relates to a model assumption. In the paper, we present data supporting that in reality the procurement policy of Global Fund recipients in terms of order timing and quantity is indeed predominantly determined by the disbursement schedule of grants. In addition, we have also edited that part of the text to clarify that the word highlighted by the referee qualifies a model assumption.

All the following other minor points are typos or formatting issues which have been addressed in the revised version of the paper:

* page 2, line 48-49: medicine related (not medicines)
* page 4, line 53-54: 'and data quality problems' (remove the 'etc', insert an 'and'. the words 'for example' appeared earlier in the sentence).
* page 13-14:equations (Input-PLT-IR-2) and (Input-b-SF-14) should be displayed equations - right now the text is too tight around them.
* First sentence of 3.2: remove 'next' as well as the second and third usages of the word 'we'.
* Page 16, line 50-51: spacing issue for the 67ppm, similar on page 17 line 26-27
* page 19: too much blank space after 'resp.' in three places - if using latex, need to use '.\ '
* page 24, line 5-6: remove 'in principle'

Reviewer 2

Comments to the Author
Comments to Authors
Title: "National Drug Stockout Risks and the Global Fund Disbursement Process for Procurement"
Production and Operations Management

In the revised version of the paper the authors have considered the comments raised by the reviewers, improving significantly the manuscript's quality. However, there are still some issues that need to be addressed. Below I provide a list, intending to help the authors in their effort.

Thank you very much for the time and efforts spent reviewing our revised paper and for highlighting related improvement opportunities. In the following, we provide a detailed response to each of your comments, including description of how we modified the paper accordingly whenever applicable.

Main text

1) The revised version of the Introduction is very well structured. At the end of this Section the authors discuss the three research questions. Research question 1 (p. 3) raises the importance of "grant performance monitoring frequency" used by the Global Fund. I would suggest the authors to discuss "grant performance monitoring frequency" and what it means before using it in the research question to improve the clarity for the reader.

Thank you for this suggestion. In the revised version of the paper, we now define the grant performance monitoring frequency within the statement first research question. In addition, the second paragraph on p.2 provides a description of performance-based financing which allows the reader to understand the subsequent reference to grant performance monitoring.

2) The authors refer in the text to Figure 2 before referring to Figure 1. Moreover, the reference to Figure 2 appears in p. 9, while it is only presented in p. 15.

Thank you for pointing out this oversight. We have now put the Figures in the correct order so that the methodological diagram (now Figure 1) appears on p.10, directly after its reference in the text.

3) Reading the process modifications presented in Subsection 3.1.4 I was wondering if the authors could discuss how often these three modifications have been used by the Global Fund already in the timeframe under study or if GF is only thinking of using them in the future.

These process modifications were discussed and considered by the Global Fund during the period of study, which motivated us to analyze them, but none of them were implemented until late 2012 so that they didn't affect the baseline data that we worked with. In the revised version of the paper we now clarify this in Section §1.2 to avoid any related confusion.

4) Figure 2 presents that "Global Fund grants disbursements data (1791 disbursements) Historical DIT". 1791 is different from the 2068 disbursement data mentioned in p. 15. The authors need to explain the reason for this difference in the main text. In general the authors should explain all the numbers that are presented in the text; this is for example true also for the 429 grants discussed in p. 18 (last paragraph), or for the 541 data instances (p. 20, 3rd paragraph), for the 1658 DIT observations in appendix A.2 (which is different from Figure 2), etc.

Thank you for pointing out these issues. In the revised paper, Section §3 introduces the complete datasets and we have edited Figure 1 to make it consistent with the total number of observations in each dataset. The related data exclusions performed as part of the estimation processes are described and justified in the relevant sections of the online supplement. Additionally, we clarified on p. 18 that 429 grants with at least three grants starting before 1 January 2007 are out of the total 461. We also clarified on p.20 that 541 data instances are the (i, j, l, m) combinations in our data set, which is larger than the number of grants since some HIV and malaria grants can be used for the purchase of both treatment and prevention drugs.

5) In p. 17 the authors surprisingly state that "the present work does not aim theoretical contributions". Since this statement cannot be true for any academic paper, I would suggest them to rephrase the text.

Thank you for this suggestion. We agree, and have now replaced the adjective "theoretical" with "methodological" in the revised version of the paper.

6) Since l stands for reporting frequency and m for product type, I would suggest the authors to change (p. 20) "They involved simulation runs for every (i; j; l; m) combination of principal recipient, country, product type and reporting frequency in our dataset" to "They involved simulation runs for every (i; j; l; m) combination of principal recipient, country, reporting frequency and product type in our dataset". The same comment holds also for p. 25.

Thank you for pointing out these inconsistencies, which are now corrected in the revised paper.

7) The authors use both "unmet" and "lost" demand throughout the text. I would suggest them to use only one of the two.

Thank you for pointing out this inconsistency – in the revised paper version, we have replaced all instances of the word "unmet" with "lost" to avoid any confusion.

8) Please change (p. 31) "systems dynamics" to "system dynamics".

We have corrected this typo in the revised version of the paper.

Appendices

9) In Appendix A.1 (p. 1) the text reads "We analysed all 3027 procurement orders initiated by principal recipients in 53 African countries between January 2008 and March 2012", while in p. 14 in the main text the period studied is between January 2002 and March 2012.

Thank you for pointing out this typo, which we have corrected in the revised paper (replacing "January 2008" with "January 2002").

10) In Appendix A.2 please elaborate more on the differences between the four regression specifications.

As a result of your comment, we have explicitly listed the variables in each of the regression specifications in the notes for Table A.4. For consistency, we have also done the same for Table A.1.

11) Figure B1 is not clear to me. For example the authors discuss that "Within those, the Global Fund's support as a fraction of total donor funding ranges between 37% in Lesotho to 94% in Gambia". Doesn't this mean that in the case of Gambia 94% of funding comes from GF and 6% from other sources? Shouldn't then the total presented in Figure B1 always be equal to 100%? Furthermore, I cannot find Lesotho in the figure.

Thank you for these comments - we recognize that the use of different chart formats across HIV, Malaria and Tuberculosis in Figure B1 in the previous version was confusing. In the revised paper, we have now edited the graph showing HIV expenditures in order to make its format and data shown (proportion of total donor funding from GF) consistent with the other two graphs for Malaria and TB.

12) What do "ITNs" (p. 10) stand for?

Thank you for pointing out this undefined acronym - ITNs stands for "insecticide-treated nets", which we now use instead in the revised paper version.

13) In the appendices there is no reference to Table A3 in the text. The same holds also for Table A6, Table C1 and Figure C3.

Thank you for this note. In the revised paper, we added information about Table A3 (p.4 of the online supplement), Table A6 (p.8) and Figure C1 (p.16). We removed Figure C3, since we discuss the effect of cash buffer in the main text (p.26-27 and Figure 3).

14) There are still some grammar mistakes and typos throughout the main text and the appendices. For example in p. 20 the main text reads "in the following we discuss the results of extensive simulation experiments". In the appendices in p. 2 the text reads "These result suggest that", p. 10 "Within those, the Global Fund's support as a fraction of total donor funding ranges between 37% in Lesotho to 94% in Gambia. Matching this data with the total HIV expenditure from domestic sources in 2012" and "indicating that the Fund is responsible for most of the TB procurement in Africa. While country-specific TB-related domestic health expenditures in Africa are not available," etc. Another proofreading is strongly recommended.

Following this comment, we have proofread the paper multiple times and have fixed various typos, including the ones mentioned. We have also separated the cumulative probability distributions of PLT and DIT in a separate subsection in the online supplement (D.1) to avoid confusion.

Reviewer: 3

Comments to the Author
I would like to commend the authors for having responded adequately to the concerns raised in the previous round. Thus, I recommend that the paper be accepted.

Thank you for your positive comments. We appreciate the time and effort you spent on this review, including your feedback in the earlier round, and are glad that you believe our paper should be published.