A Hierarchical Bayesian Model of Pitch Framing

Sameer K. Deshpande and Abraham J. Wyner

The Wharton School
University of Pennsylvania

9 September 2017

Abstract

1 Introduction

Fig 1: B[illegible] ([illegible]) [illegible]
[illegible] K[illegible] ([illegible]) [illegible] S
[illegible] h[illegible]://[illegible]/

B[illegible] *strike zone*, [illegible] A[illegible] f
[illegible] t a[illegible]
K[illegible] As a[illegible],
[illegible] C[illegible] C[illegible]
[illegible]
[illegible] (s e.g. P[illegible] et al., 2011; [illegible] et al., 2015), [illegible] (s e.g. Ki[illegible] and K[illegible], 2014; M[illegible], 2014), [illegible] (C[illegible] et al., 2016). D[illegible]
[illegible] C[illegible]
A[illegible] J[illegible] C[illegible] C[illegible]
[illegible] ([illegible], 2015).

[illegible] 2008, t[illegible]
[illegible] (s e.g. L[illegible], 2013; P[illegible], 2014; [illegible], 2015) [illegible] (s e.g. D[illegible], 2014; H[illegible], 2014) [illegible]

[illegible] A[illegible] (2014),

[illegible] J[illegible] (2015) [illegible] B[illegible]

P[illegible] 25 [illegible]

[illegible] B[illegible] 10 [illegible]

[illegible] (C[illegible] 2008), [illegible]

[illegible] 2 to 3 [illegible]

[illegible] D[illegible]

[illegible] (2015) [illegible]

J[illegible] L[illegible] 2 [illegible] 2014 [illegible] $14M,

[illegible] 17th [illegible]"

[illegible] Fig. 1, C[illegible] K[illegible]

[illegible] C[illegible] H[illegible] Fig. 1, [illegible]

[illegible]

to C[illegible]. F[illegible] K[illegible] F[illegible]

[illegible] 1 – 0 [illegible] 1 – 1 [illegible] W[illegible]

[illegible] I[illegible]

K[illegible] A[illegible]

[illegible] C[illegible] K[illegible] O[illegible]

[illegible] C[illegible]

[illegible] C[illegible] C[illegible]

[illegible]

A[illegible]

E[illegible] fit

[illegible] 6

[illegible] (2008) [illegible] F[illegible]

[illegible] (2007).

[illegible] (2008) [illegible]

[illegible] 0.16 [illegible] M[illegible]

(2015)

(2011) fit a

(2008)'s

(2015).

fit

fit a

linear

0 or 1 (e.g.

(2015)'s (2011)'s

[illegible]

B[illegible]. F[illegible] $y = 1$ [illegible] $y = 0$ [illegible]. Let $\mathbf{b}, \mathbf{ca}, \mathbf{co}, \mathbf{p}$ and $\mathbf{u}$ [illegible]. F[illegible] x and z [illegible]. T[illegible] u [illegible] $\Theta^{u,B}, \Theta^{u,CA}, \Theta^{u,P}$, and $\Theta^{u,CO}$ [illegible] *partial effect* [illegible] u's [illegible]. F[illegible] u, [illegible] $f^u(x, z)$, [illegible]. A[illegible]

$$\log\left(\frac{\mathbb{P}(y=1)}{\mathbb{P}(y=0)}\right) = \Theta^{\mathbf{u},B}_{\mathbf{b}} + \Theta^{\mathbf{u},CA}_{\mathbf{ca}} + \Theta^{\mathbf{u},P}_{\mathbf{p}} + \Theta^{\mathbf{u},CO}_{\mathbf{co}} + f^{\mathbf{u}}(x, z) \tag{1}$$

[illegible] PITCHf/x [illegible] 2.1, [illegible]. In [illegible] 2.3, [illegible] 1 [illegible]. Al[illegible] (D[illegible], 2014; H[illegible], 2014; [illegible], 2015). [illegible] 3.1 [illegible] 4 [illegible]

2 Data and Model

[illegible] (2010) [illegible] (2014), [illegible]

2.1 PITCHf/x Data

In 2006, [illegible] PITCHf/x [illegible] 27 [illegible] PITCHf/x [illegible] 27 [illegible] ([illegible], 2014). [illegible] MLB [illegible] ([illegible], 2010). In [illegible] MLB [illegible] [illegible] (e.g. [illegible]), [illegible] (e.g. [illegible]), [illegible] (e.g. [illegible]). [illegible] PITCHf/x [illegible] MLB [illegible] API.

[illegible] 2014 [illegible] PITCHf/x [illegible] 2011 [illegible] 2015 [illegible]. In [illegible] 2.2, [illegible] 2011–2013 [illegible] $f^u(x,z)$ [illegible] 1. [illegible] 2014 [illegible] 3.1, [illegible] 2015. In [illegible] 2014 [illegible] 701,490 [illegible] 355,293 (50.65%) [illegible] *taken* (i.e. [illegible]) [illegible] 124,642 (35.08%) [illegible]

Moreover [illegible]

PITCHf/x [illegible] $N = 308,388$

In [illegible]

$n_U = 93$ [illegible] $n_B = 1010$ [illegible] $n_C = 101$ [illegible] $n_P = 719$ [illegible]

2.2 Adjusting for Pitch Location

[illegible]

(x, z) [illegible] PITCHf/x [illegible] $f^u(x, z) = \theta_x^u x + \theta_z^u z$, [illegible]

θ_x^u and θ_z^u [illegible]

A [illegible]

[illegible]

Fig. 2, [illegible]

"1" [illegible]

2011 – 2013 [illegible] A [illegible] Fig. 2

[illegible]

PITCHf/x [illegible] If [illegible]

[illegible]

(i.e. [illegible])

(i.e. [illegible]) [illegible] 6

12 [illegible]

12 [illegible]

[illegible] 1 [illegible]

[illegible]0 [illegible]

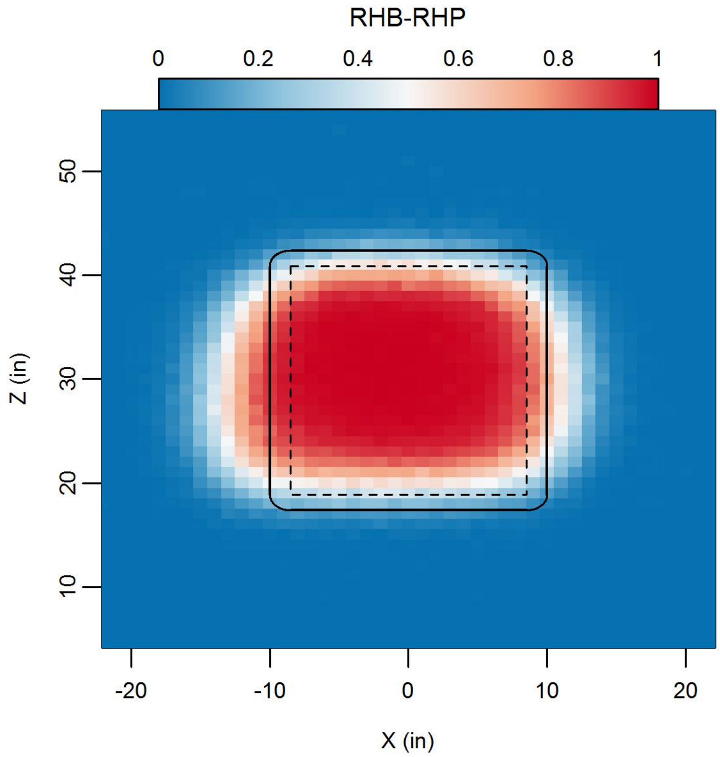


Fig 2: H[illegible]
2011 –2013. [illegible] 2014 [illegible]
If [illegible]
[illegible] = 100% [illegible] =
50%, [illegible] = 0%.

[illegible]

[illegible]

[illegible] J[illegible] (2015), [illegible]

[illegible]

Fig 3 [illegible]

[illegible] O[illegible]

[illegible]

[illegible]

[illegible]

[illegible] m

[illegible] In [illegible], in [illegible]

[illegible] m

[illegible] in Fig [illegible] 3 [illegible]

a [illegible]

[illegible]

I [illegible] M[illegible] (2014), [illegible] it [illegible]

[illegible] 2011 –2013, [illegible]

[illegible] GAMS" [illegible] GAMS" [illegible]

[illegible] a [illegible] Fig [illegible] 4 [illegible] GAM [illegible]

[illegible] GAMs

[illegible] Fo

[illegible] GAMs [illegible]

[illegible] m

[illegible] Fo [illegible] 2014 [illegible]

[illegible] GAM [illegible]

[illegible]

[illegible]

2.3 Bayesian Logistic Regression Models

B [illegible] 93 [illegible] $u_1, \ldots, u_{93}$. C [illegible] i^{th} [illegible]

[illegible] $y_i = 1$ [illegible] $y_i = 0$ [illegible] Le [illegible] h_i [illegible]

[illegible] $\mathbf{LO}_i$ [illegible]

[illegible]

[illegible] GAM [illegible] x_i [illegible] z_i [illegible] PITCHf/x

[illegible] $f^u(x_i, z_i) = h_i^\top \Theta_0^u + \mathbf{LO}_i^\top \Theta_{LO}^u$, [illegible] Θ_{LO}^u [illegible]

[illegible] Θ_0^u [illegible]

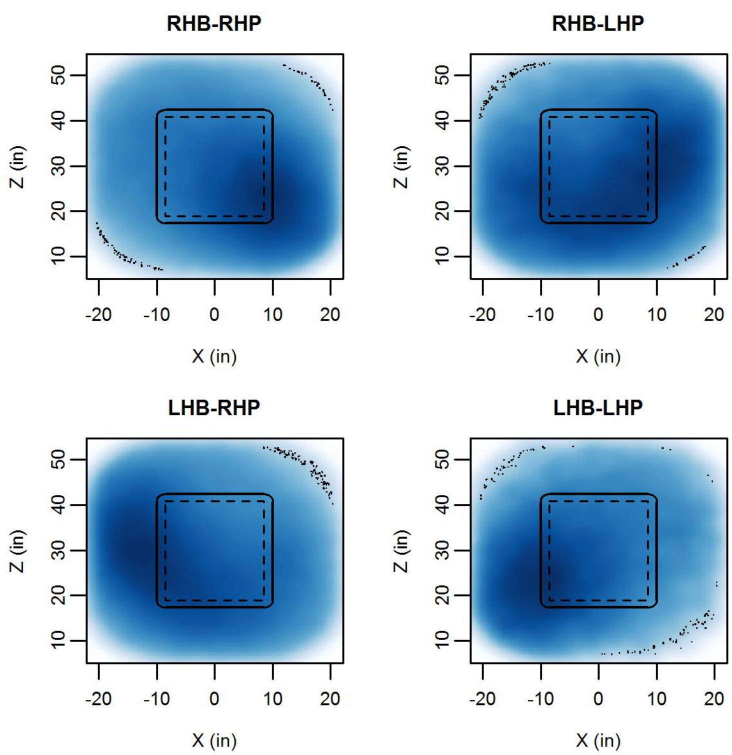


Fig 3: K[illegible] F[illegible]
[illegible]
[illegible] D[illegible]

[illegible] F[illegible] $u(i)$ [illegible]

[illegible]

[illegible] GAM [illegible]

[illegible].

F[illegible] $\mathbf{CO}_i, \mathbf{CA}_i, \mathbf{P}_i$ [illegible] $\mathbf{B}_i$ [illegible]

[illegible] $\Theta_{CO}^{u}, \Theta_{CA}^{u}, \Theta_{P}^{u}$, [illegible] Θ_{B}^{u} [illegible]

[illegible] u. F[illegible] B[illegible]

P[illegible] 0 –0, [illegible] E[illegible] 1 [illegible]

$$\log\left(\frac{\mathbb{P}(y_i = 1)}{\mathbb{P}(y_i = 0)}\right) = h_i^\top \Theta_0^{u(i)} + \mathbf{LO}_i^\top \Theta_{LO}^{u(i)} + \mathbf{CO}_i^\top \Theta_{CO}^{u(i)} + \mathbf{CA}_i^\top \Theta_{CA}^{u(i)} + \mathbf{P}_i^\top \Theta_{P}^{u(i)} + \mathbf{B}_i^\top \Theta_{B}^{u(i)}$$

[illegible]

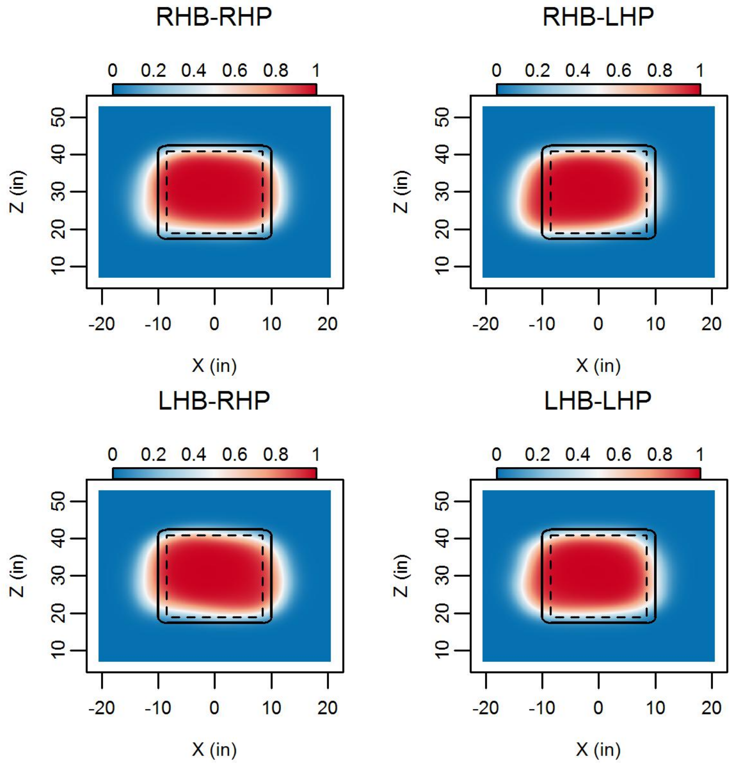


Fig 4: GAM [illegible] R = 100% [illegible] = 50%, [illegible] = 0%.

[illegible] (ie [illegible] $\Theta^u_{CO}, \Theta^u_{CA}, \Theta^u_P$, [illegible] Θ^u_B [illegible]

[illegible] Md [illegible]

[illegible]

a priori,

$$\Theta_0^{u_1}, \ldots, \Theta_0^{u_{93}} | \Theta_0 \sim N\left(\Theta_0, \tau_0^2 I_4\right)$$

$$\Theta_{LO}^{u_1}, \ldots, \Theta_{LO}^{u_{93}} | \Theta_{LO} \sim N\left(\Theta_{LO}, \tau_{LO}^2 I_4\right)$$

$$\Theta_0 | \sigma_0^2 \sim N\left(0_4, \sigma_0^2 I_4\right)$$

$$\Theta_{LO} | \sigma_{LO}^2 \sim N\left(\mu_{LO}, \sigma_{LO}^2 I_4\right)$$

[illegible] μ_{LO} [illegible] GAM [illegible]

[illegible] In [illegible] Model 1 [illegible] GAM [illegible] τ_0^2 and τ_{LO}^2 [illegible] Θ_0 and Θ_{LO} [illegible] By [illegible] Θ_0 and Θ_{LO}, [illegible] Θ_0^u's and Θ_{LO}^u's [illegible] *a priori* and *a posteriori*. In [illegible] u's [illegible] Θ_0^u [illegible] τ_0^2 and τ_{LO}^2. F[illegible] σ_0^2 and σ_{LO}^2 [illegible] Θ_0^u and Θ_{LO}^u [illegible]

[illegible] Model 1 [illegible] Model 1. In Model 2, [illegible] Model 2 [illegible] Θ_{CO}^u's [illegible] Θ_{CO} [illegible] Θ_{CA}^u's [illegible] Θ_{CA}. [illegible] Model 3 [illegible] Model 2 [illegible] *A priori*, [illegible]

$$\Theta_{CO}|\sigma^2_{CO} \sim N\left(0_{11}, \sigma^2_{CO} I_{11}\right)$$

[illegible] Θ_{CA}, Θ_P and Θ_B, [illegible] Θ_0^u's and Θ_{LO}^u's [illegible] Model 1, Model 2 and 3 [illegible] E[illegible] 1 [illegible] Model 4, [illegible] Model 2 [illegible] and Model 5, [illegible]

[illegible] Eq. 1 [illegible]

$$\Theta_{CO}^{u_1}, \cdots, \Theta_{CO}^{u_{93}}|\Theta_{CO}, \tau_{CO}^2 \sim N\left(\Theta_{CO}, \tau_{CO}^2 I_{11}\right)$$

$$\Theta_{CO}^{u}|\sigma_{CO}^2 \sim N\left(0_{11}, \sigma_{CO}^2 I_{11}\right)$$

[illegible] $\Theta^u_{CA}, \Theta^u_B, \Theta^u_P$ in Model 4 and 5. [illegible]

[illegible] Ga(3,3) [illegible]

$\sigma_0^2, \sigma_{LO}^2, \sigma_{CO}^2, \sigma_{CA}^2, \sigma_P^2$ [illegible] σ_B^2.

It [illegible] $\tau_0^2, \tau_{LO}^2, \tau_{CO}^2, \tau_{CA}^2, \tau_P^2$ [illegible] τ_B^2 [illegible]

[illegible] F[illegible]

[illegible] 0.25 i[illegible]

[illegible] GAM [illegible] 50%

[illegible] A[illegible] Model 2, [illegible]

[illegible] $N\left(0, 2(\tau_0^2 + \tau_{CO}^2 + \tau_{CA}^2)\right)$ [illegible] *a priori.* [illegible] $\tau_0^2 = \tau_{CO}^2 = \tau_{CA}^2 =$ 0.25 [illegible] 10% [illegible]

75% [illegible] 25% [illegible] F[illegible]

$\tau_{LO}^2 = \tau_B^2 = \tau_P^2 = 0.25$ [illegible]

3 Model Performance and Comparison

3.1 Predictive Performance

[illegible] (C[illegible], 2016) [illegible] MCMC [illegible] A1

[illegible] R [illegible] 3.3.2 [illegible] MCMC [illegible]

[illegible] 2.14.1 [illegible] F[illegible]

[illegible] 2000 [illegible] $\hat{R}$ [illegible]

[illegible] 1.1, [illegible]

[illegible] 1000. For Model 1 and 2, [illegible]

[illegible] 4000 [illegible] Models 3, 4, and 5, [illegible] 6000 [illegible] (Model 1) to 50 [illegible] (Model 5).

[illegible]

[illegible] 2014 [illegible] Table 1 [illegible] Model 1 – 5 [illegible] 2014 [illegible] hGAM [illegible] Region 1 [illegible] 1.45 [illegible] Section 2.1. [illegible] 1.45 [illegible] Region 1 [illegible] Region 2 [illegible] 1.45 and 2.9 [illegible] 0.5 [illegible]

Table 1: [illegible]

	# Parameters	Model 1 744	Model2 855	Model 3 2582	Model 4 11,067	Model 5 171,168	hGAM – –
Overall	MISS	0.103	0.100	0.099	0.096	**0.0856**	0.105
	MSE	0.073	0.071	0.069	0.068	**0.061**	0.074
Region 1	MISS	0.248	0.236	0.232	0.225	**0.195**	0.258
	MSE	0.163	0.156	0.153	0.150	**0.133**	0.168
Region 2	MISS	0.214	0.209	0.205	0.203	**0.184**	0.215
	MSE	0.153	0.149	0.146	0.144	**0.129**	0.156

[illegible] Model 1 – 5 [illegible] GAMs [illegible] Region 1 and 2. [illegible] GAMs [illegible] 2011 – 2013 [illegible] 2014 [illegible] Model 1 [illegible] As [illegible] (Model 2), [illegible] (Model 3) [illegible] Model 5 [illegible] Model 5 [illegible]

O[illegible] 2

[illegible] 2015 [illegible]

[illegible], [illegible] 2014 [illegible]

[illegible] 2: O[illegible]

	# Parameters	Model 1 744	Model2 855	Model 3 2582	Model 4 11,067	Model 5 171,168	hGAM – –
Overall	MISS	0.107	0.105	**0.105**	0.106	0.106	0.109
	MSE	0.075	0.074	**0.074**	0.075	0.074	0.076
Region 1	MISS	0.256	0.245	**0.244**	0.248	0.246	0.267
	MSE	0.167	0.162	**0.161**	0.163	0.162	0.173
Region 2	MISS	0.236	0.232	**0.231**	0.233	0.234	0.237
	MSE	0.169	0.166	**0.165**	0.166	0.165	0.170

N[illegible] Model 3 [illegible] Region 1 [illegible] 2. [illegible] Models 4 and 5 [illegible] O[illegible] 2015 [illegible] [illegible] (2014), M[illegible] (2017a), and M[illegible] (2017b) [illegible] M[illegible] L[illegible] B[illegible] 2009. In Appendix A, [illegible] fit Models 1 –5 [illegible] 90% [illegible] 2014 [illegible] 10%, [illegible] Model 3's [illegible]

Model 3's [illegible] Models 1 and 5 [illegible]

3.2 Full Posterior Analysis

[illegible] Model 3 [illegible] Figure 5 [illegible] 6 [illegible] 10 [illegible] 10 [illegible] 10 [illegible]

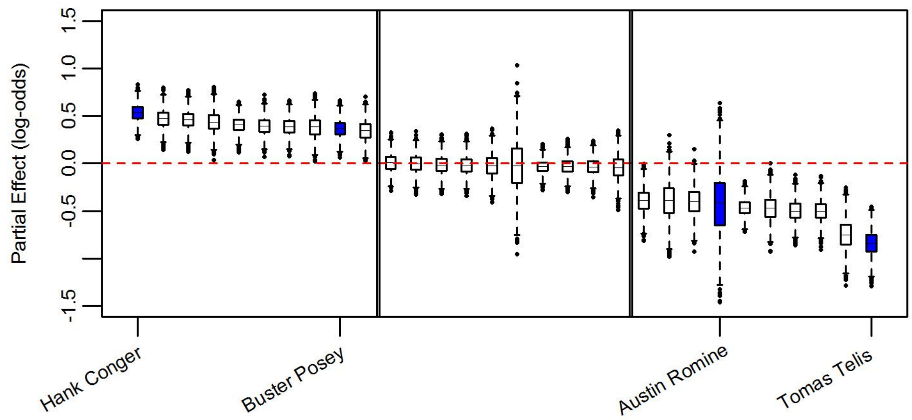


Fig 5: [illegible] 30 [illegible]

[illegible] Hank Conger [illegible] Buster Posey [illegible]

[illegible]

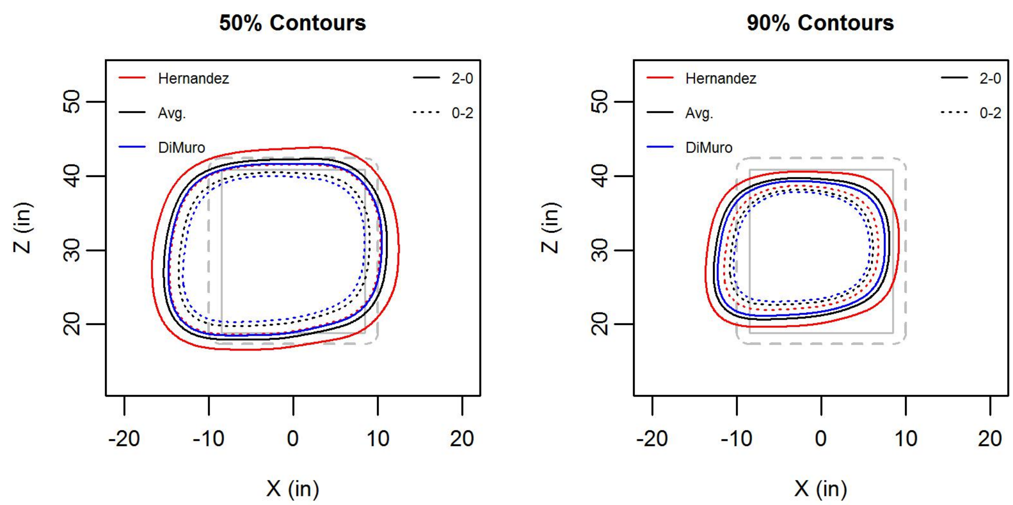
[illegible] [−1.5, 1.5] [illegible] 20% [illegible]

[illegible] 3 [illegible] 18.24% [illegible] 81.76%. [illegible] [−1.5, 1.5] [illegible]

[illegible] Appendix B.

[illegible]

[illegible] P[illegible] B[illegible] P[illegible] Fig[illegible] 6 [illegible] 50% [illegible] 90% [illegible] A[illegible] H[illegible] M[illegible] DiMu[illegible] 2 –0 [illegible] 0 –2 [illegible] N[illegible] [illegible] P[illegible] a[illegible]



Fig[illegible] 6: 50% [illegible] 90% [illegible] B[illegible] P[illegible] P[illegible]

A[illegible] H[illegible] DiM[illegible] F[illegible] 2 –0 [illegible] H[illegible] 50% [illegible] 4.37 sq f, [illegible] 3.87 sq f, [illegible] DiM[illegible] 3.53 sq f. [illegible] 0 –2 [illegible] 0 –2 [illegible] 2 –0 [illegible] E[illegible] 50% [illegible] *inside* [illegible] A[illegible] *outside* [illegible] H[illegible] DiM[illegible] E[illegible] 90% [illegible] 2 –0 [illegible]

DM, H, DM

DM

90% 0 −2

A

DM

90%

F

90% 0 −2

S

90% 50%,

Fig. 6 H

DM H 42.67% (1624

2182) DM 39.92% (1220 1836).

On 2 −0 H 51.44% (71 67) DM

48.31% (57 67).

4 Impact of framing

F S

2.3, $\mathbf{h}$

$\mathbf{lo}$

GAM. L $\mathbf{b}, \mathbf{ca}, \mathbf{co}, \mathbf{p}$ $\mathbf{u}$

F B P ca_0. F

$\xi = (\mathbf{u}, \mathbf{co}, \mathbf{lo}, \mathbf{b}, \mathbf{p}, \mathbf{h})$

[illegible] ($\mathbf{ca}, \xi)$. F[illegible] ca, [illegible] $\mathcal{P}_{ca}$ [illegible] ca:

$$\mathcal{P}_{ca} = \{(\mathbf{ca}, \xi) : \mathbf{ca} = ca\} .$$

F[illegible] $TAKEN$ [illegible] $CALL \in \{Ball, Strike\}$ [illegible] W[illegible] S, [illegible] ($\mathbf{ca}, \xi)$, [illegible] A [illegible] S [illegible]

$$\mathbb{E}[S|\mathbf{ca}, \xi, TAKEN] = \sum_{CALL} \mathbb{E}[S|COUNT, TAKEN, CALL]\mathbb{P}\left(CALL|\mathbf{ca}, \xi, TAKEN\right)$$

[illegible] ca, [illegible] B[illegible] P[illegible] I[illegible] $E\left[S|\mathbf{ca} = ca_0, \xi, TAKEN, CALL\right]$. W[illegible]

$$\mathbb{E}\left[S|\mathbf{ca} = ca, \xi, TAKEN, CALL\right] - \mathbb{E}\left[S|\mathbf{ca} = ca_0, \xi, TAKEN, CALL\right]$$

[illegible] (ie [illegible]) [illegible] ca's [illegible] A [illegible]

$$f(ca, \xi) = (\mathbb{P}(Strike|\mathbf{ca} = ca, \xi, TAKEN) - \mathbb{P}(Strike|\mathbf{ca} = ca_0, \xi, TAKEN)) \times \rho(COUNT),$$

[illegible]

$$\rho(COUNT) = \mathbb{E}[S|COUNT, TAKEN, Ball] - \mathbb{E}[S|COUNT, TAKEN, Strike].$$

[illegible] *ca*'s *framing effect*:

[illegible]

[illegible] Model 3. [illegible] ρ [illegible]

[illegible]

[illegible] ρ, [illegible] 182,405 0–1 [illegible] (140,667 [illegible] 41,738 [illegible] 2011 [illegible] 2014. [illegible] 0.322 [illegible] 0–1 [illegible] 0.265 [illegible] 0–1 [illegible] 0–1 [illegible] 0.057 [illegible] 3 [illegible] ρ. [illegible] 2011 [illegible] 2014. [illegible] 3–2 [illegible] 2.1% [illegible] 3–2 [illegible] (1963), [illegible] (2010) [illegible] (1985) [illegible] (2015) [illegible]

[illegible] 3 [illegible] 0.11 [illegible] 0.14 [illegible] (2015) [illegible] 0.161 [illegible] (2008). [illegible]

[illegible]: E[illegible]

Count	Ball	Strike	Value of Called Strike ρ	Proportion
$0-0$	0.367 (0.002)	0.305 (0.002)	0.062 (0.002)	36.2%
$0-1$	0.322 (0.002)	0.265 (0.004)	0.057 (0.004)	12.5%
$0-2$	0.276 (0.003)	0.178 (0.007)	0.098 (0.008)	5.5 %
$1-0$	0.427 (0.003)	0.324 (0.003)	0.103 (0.005)	11.5%
$1-1$	0.364 (0.003)	0.280 (0.004)	0.084 (0.005)	8.8%
$1-2$	0.302 (0.003)	0.162 (0.006)	0.140 (0.006)	6.9%
$2-0$	0.571 (0.007)	0.370 (0.006)	0.201 (0.009)	3.9%
$2-1$	0.468 (0.005)	0.309 (0.006)	0.159 (0.008)	4.0%
$2-2$	0.383 (0.004)	0.165 (0.006)	0.218 (0.007)	4.8%
$3-0$	0.786 (0.013)	0.481 (0.008)	0.305 (0.015)	1.9%
$3-1$	0.730 (0.010)	0.403 (0.009)	0.327 (0.014)	1.8%
$3-2$	0.706 (0.008)	0.166 (0.008)	0.540 (0.011)	2.1%

[illegible] ρ [illegible] $f(ca,\xi)$ [illegible] ca's [illegible]

$$RS(ca) = \sum_{(\mathbf{ca},\xi)\in\mathcal{P}_{ca}} f(ca,\xi).$$

[illegible] (2015) [illegible] (2015) [illegible] (2015). [illegible] 4 [illegible] 10 [illegible] 95% [illegible] (2015)'s [illegible]

[illegible] 10 [illegible] 4 [illegible] 10 [illegible]

ble4: pdaBt10 higtohpofRTcBP
inJgeta (2015)'shtopahBlP plsw

Rank	Catcher	Runs Saved (SD)	95% Interval	N	BP
1	Miguel Montero	25.71 (5.03)	[15.61, 35.09]	8086	11.2 (8172)
2	Mike Zunino	22.72 (5.17)	[12.56, 32.31]	7615	20.4 (7457)
3	Jonathan Lucroy	19.56 (5.69)	[8.16, 30.49]	8398	16.4 (8241)
4	Hank Conger	19.34 (3.24)	[12.93, 25.65]	4743	23.8 (4768)
5	Rene Rivera	18.81 (3.69)	[11.63, 25.89]	5091	22.5 (5182)
6	Buster Posey	17.01 (4.14)	[8.79, 25.01]	6385	23.6 (6190)
7	Russell Martin	14.35 (4.41)	[5.85, 22.77]	6388	14.9 (6502)
8	Brian McCann	14.01 (3.95)	[6.18, 21.66]	6335	9.7 (6471)
9	Yasmani Grandal	12.88 (2.98)	[7.18, 18.69]	4248	14.5 (4363)
10	Jason Castro	12.61 (4.43)	[3.80, 21.08]	7065	11.5 (7261)
92	Josmil Pinto	-6.49 (1.41)	[-9.32, -3.76]	1748	-6.9 (1721)
93	Welington Castillo	-6.70 (4.28)	[-15.19, 1.78]	6667	-15.6 (6661)
94	Chris Iannetta	-7.50 (4.46)	[-16.18, 1.08]	6493	-7.3 (6527)
95	John Jaso	-7.76 (2.41)	[-12.50, -3.07]	3172	-11.3 (2879)
96	Anthony Recker	-8.37 (2.33)	[-13.29, -3.93]	2935	-13 (3102)
97	Gerald Laird	-8.68 (1.87)	[-12.29, -4.99]	2378	-9.6 (2616)
98	A. J. Ellis	-12.90 (3.79)	[-20.10, -5.38]	5476	-12.3 (5345)
99	Kurt Suzuki	-17.67 (4.25)	-[26.07, -9.35]	6811	-19.5 (7110)
100	Dioner Navarro	-18.81 (4.68)	[-28.00, -9.40]	6659	-19.8 (6877)
101	Jarrod Saltalamacchia	-23.98 (4.35)	[-32.76, -15.87]	6498	-34 (6764)

haMgMofaste25.71 sngdehsa T

is dehpyubnfa8,086 dehpdsh

tolehnepahd25.71 n, ogUcsf

gipfeJgeta (2015)'sehsFgtof a

ihthhfaks dofpyI el

imsa (eg MoheR, yhbptsn

ofsnshishs, erepb(eg MkZn JnLy).

N sbhdwpbihsbhoa

ihosba0.94. Mpfahsbhdep

eisl wfglhshsbha95% ip

hdha

4.1 Catcher Aggregate Framing Effect

L[illegible] 4, [illegible] M[illegible] A[illegible] [illegible] 3086 [illegible] C[illegible] 4743. H[illegible]

A [illegible] M[illegible] C[illegible] I[illegible] 4 [illegible]

[illegible] *integrate* $f(ca, \xi)$ [illegible] ξ [illegible] $f(ca, \xi)$ [illegible] $\mathcal{P}_{ca}$. [illegible] (CAFE) [illegible] (2009). [illegible] W [illegible] $f(ca, \xi)$ [illegible] ξ [illegible] ca's [illegible] CAFE [illegible]

$$CAFE(ca) = 4000 \times \frac{1}{N} \sum_{\xi} f(ca, \xi), \tag{2}$$

[illegible] 2 [illegible] ca [illegible] 4000 [illegible] CAFE [illegible] 4000 [illegible] 25 [illegible] 3,992. Of [illegible]

O[illegible] Θ^{u}'s [illegible]

[illegible] CAFE. [illegible] 5 [illegible] 10 [illegible]

[illegible] CAFE [illegible], [illegible] 95%

[illegible] CAFE [illegible] A[illegible] 95% [illegible]n

[illegible] CAFE.

[illegible] 5: [illegible] B[illegible] 10 [illegible] $CAFE$.

Rank	Catcher	Mean (SD)	95% Interval	95% Rank Interval
1.	Hank Conger	16.20 (2.72)	[10.84, 21.50]	[1, 11]
2.	Christian Vazquez	14.33 (2.94)	[8.26, 20.03]	[1, 19]
3.	Rene Rivera	14.04 (2.76)	[8.75, 19.31]	[1, 18]
4.	Martin Maldonado	13.24 (3.33)	[6.73, 19.68]	[1, 24]
5.	Miguel Montero	12.36 (2.42)	[7.50, 16.90]	[2, 22]
6.	Yasmani Grandal	11.90 (2.76)	[6.56, 17.29]	[2, 27]
7.	Mike Zunino	11.78 (2.69)	[6.51, 16.74]	[2, 26]
8.	Chris Stewart	11.63 (3.28)	[5.21, 18.03]	[1, 30]
9.	Buster Posey	11.16 (2.73)	[5.74, 16.51]	[2, 30]
10.	Francisco Cervelli	10.45 (3.21)	[4.06, 16.72]	[2, 36]
92.	Jordan Pacheco	-11.73 (3.80)	[-19.26, -4.30]	[68, 98]
93.	Koyie Hill	-11.79 (5.67)	[-22.48, -0.68]	[53, 100]
94.	Josh Phegley	-12.05 (4.66)	[-21.40, -3.20]	[64, 99]
95.	Austin Romine	-12.76 (9.78)	[-32.14, 5.81]	[30, 101]
96.	Jarrod Saltalamacchia	-14.00 (2.53)	[-19.11, -9.26]	[82, 99]
97.	Brett Hayes	-14.04 (4.06)	[-21.51, -5.93]	[73, 100]
98.	Gerald Laird	-14.96 (3.21)	[-21.17, -8.69]	[81, 99]
99.	Josmil Pinto	-15.04 (3.27)	[-21.57, -8.78]	[82, 100]
100.	Carlos Santana	-22.48 (4.63)	[-31.63, -13.26]	[93, 101]
101.	Tomas Telis	-25.06 (3.85)	[-32.41, -17.27]	[98, 101]

W[illegible] 4 [illegible] 5. [illegible]

[illegible], C[illegible] V[illegible] M[illegible] M[illegible], C[illegible] S[illegible] F[illegible] C[illegible]

[illegible] 13^{th}, 17^{th}, 18^{th} [illegible] 19^{th} [illegible]

[illegible] ξ [illegible]

[illegible] 2014 [illegible] 10

[illegible] 4. In[illegible] 3198 [illegible] C[illegible] 2424,

[illegible] 2370, [illegible] M[illegible] 1861.

I[illegible] H[illegible] C[illegible] M[illegible] M[illegible]

[illegible] CAFE, [illegible] 4 [illegible]

[illegible]

C[illegible] M[illegible] CAFE, [illegible] [illegible] 95% [illegible]

[illegible] 95% [illegible] C[illegible] 11th [illegible] CAFE. In[illegible] 95% [illegible] CAFE [illegible] 3 CAFE [illegible] I[illegible] 10 [illegible] K[illegible] H[illegible] A[illegible] R[illegible] T[illegible] H[illegible] 409 [illegible] R[illegible] 61. As[illegible]

4.2 Year-to-year reliability of CAFE

[illegible] CAFE [illegible] W[illegible] t[illegible] 2012 t[illegible] 2015 [illegible] F[illegible] [illegible] W[illegible] GAM [illegible] 2012 [illegible] GAM [illegible] 2011 [illegible] 2015 [illegible] GAM [illegible] 2011 t[illegible] 2014. [illegible] CAFE, [illegible] 3 [illegible] 56 [illegible] 5 [illegible] CAFE [illegible]

[illegible] 5: C[illegible] CAFE [illegible]

	2012	2013	2014	2015
2012	1.00	0.70	0.56	0.41
2013	0.70	1.00	0.71	0.61
2014	0.56	0.71	1.00	0.58
2015	0.41	0.61	0.58	1.00

I[illegible] CAFE [illegible] I[illegible] 0.5 – 0.7 [illegible] CAFE [illegible]

[illegible] 2012 CAFE [illegible] 2013 CAFE [illegible] 2013 CAFE [illegible] 2014 CAFE [illegible] 0.7, [illegible] 2014 CAFE [illegible] 2015 CAFE [illegible] 0.58. [illegible] 2014 [illegible] 2015 [illegible] O [illegible] (2016).

5 Discussion

[illegible] 2014. F [illegible] (2015), [illegible] R O [illegible]

[illegible]

[illegible] CAFE [illegible] CAFE [illegible]

CAFE

CAFE

CAFE

D

CAFE, 2014 2015.

2014 O

O M L B

CAFE

F 95% J L [8.16, 30.49].

10 $7M (C, 2014; P, 2013),

L $5 .7M $21 .34M. In

CAFE,

CAFE

A M M

H C J

M C

5.1 Extensions

CAFE

O

F

(2015)

In

On

(e.g.

10

E

M

Chen et al. (2016)

k

In M

CAFE. At a $priori$

k. It

of

of

M fit

C

$f^u(x,z)$ in Eq. 1.

[illegible]

[illegible] A [illegible]

[illegible] fit a B[illegible] Chen et al. (2010)'s BART [illegible]

[illegible]

[illegible]

[illegible]

F[illegible] 2015 A[illegible] L[illegible] Fig. 1.

Fig[illegible] 2015 [illegible] E[illegible] C[illegible]

[illegible] K[illegible] (81.72% vs 62.59%). I[illegible]

[illegible] GAMs [illegible] 51.31% and 50.29%, [illegible] L[illegible]

[illegible]

[illegible] 77.58% [illegible] K[illegible] 61.29% [illegible] A[illegible]

[illegible] J[illegible] C[illegible] (4.14%) [illegible]

B[illegible] MC[illegible] (1.30%). [illegible]

[illegible]

[illegible] I[illegible]

[illegible] K[illegible] 1 –0 [illegible]

[illegible] 1 –1 [illegible] In 2015, [illegible]

[illegible] 1 –0 [illegible] 1 –1 [illegible] I[illegible] K[illegible]

[illegible] C[illegible]

[illegible] K[illegible]

[illegible] A[illegible]

[illegible] Fig. 1 [illegible]

[illegible]

References

Alb, J. (2010). [illegible] *Journal of Quantitative Analysis in Sports*, 6(4).

Alb, J. (2015). [illegible] *Journal of Sports Analytics*, 1(1):3–18.

Cam, D. (2008). [illegible] P [illegible] F [illegible] [illegible]

[illegible].

Cam, D. (2014). [illegible] C [illegible] a [illegible] in [illegible] 2014 Off-[illegible]

[illegible]2014-[illegible].

C[illegible], B., G[illegible], A., H[illegible], M., L[illegible], D., G[illegible], B., B[illegible], M., B[illegible], M. A., G[illegible], J., Li, P., [illegible] R. A. (2016). [illegible] A P[illegible] *Journal of Statistical Software*.

C[illegible], D., M[illegible], T. J., [illegible] K. (2016). D[illegible] E[illegible] [illegible] *Quarterly Journal of Economics*, 131(3):1181–1241.

C[illegible], H. A., G[illegible], E. I., [illegible] M[illegible], R. E. (2010). B[illegible] B[illegible] [illegible] *Annals of Applied Statistics*, 4(1):266–298.

D[illegible], E. (2014). [illegible] C[illegible] T M[illegible] C[illegible] A[illegible] [illegible] [illegible]2014/11/05/[illegible]

[illegible].

F[illegible], M. (2010). *What the Heck is PITCHf/x?*, [illegible] 153–158. [illegible]

H[illegible], R. (2014). H[illegible] P[illegible] F[illegible] [illegible]2014/12/12/7375383/[illegible]

[illegible]

J[illegible], S. T., [illegible] K. E., [illegible] A. J. (2009). B[illegible] A B[illegible] H[illegible] M[illegible] F[illegible] *Annals of Applied Statistics*, 3(2):491–520.

J[illegible] J., P[illegible] H., [illegible] B[illegible] D. (2015). M[illegible] B[illegible] A M[illegible] A[illegible] M[illegible] C[illegible] F[illegible] [illegible]-25514.

Kim, J. W. [illegible] K[illegible] B. G. (2014). [illegible] M[illegible] [illegible] *Management Science*, 60(11):2619–2644.

L[illegible] B. (2013). [illegible] A[illegible] P[illegible] F[illegible] [illegible]

L[illegible] G. R. (1963). An [illegible] B[illegible] *Operations Research*, 11(4):477–501.

M[illegible] M. (2011). E[illegible] C[illegible] Q[illegible] F[illegible] P[illegible] S[illegible] [illegible]

M[illegible] B. M. (2014). [illegible] P[illegible] I[illegible] A[illegible] M[illegible] E[illegible] *Managerial and Decision Economics*, 35:387–403.

M[illegible] B. M. (2017a). P[illegible] I[illegible] [illegible] *Economic Inquiry*, 55(2):1104–1118.

M[illegible] B. M. (2017b). [illegible] E[illegible] [illegible] *Labour Economics*, 46:189–199.

P[illegible] C. A., [illegible] J., [illegible] M. C., [illegible] H[illegible] D. S. (2011). [illegible] D[illegible] [illegible] *American Economic Review 101*, [illegible] 1410–1435.

P[illegible] H. (2014). [illegible] F[illegible] [illegible] /[illegible]11127248/[illegible]

P[illegible] L. (2013). How Much Does a [illegible] C[illegible] [illegible]2013/10/15/4818740/[illegible]

[illegible] J. (2014). [illegible] E[illegible] Is Out [illegible] C[illegible] [illegible]

[illegible], R. J. [illegible] S. (2015). [illegible] In *Sloan Sports Analytics Conference 2015*.

[illegible] R. (2015). [illegible] _/[illegible]12492794/[illegible]

[illegible] G. [illegible] B. (2014). MONE[illegible]: [illegible] *Annals of Applied Statistics*, 8(2):926955.

[illegible] J. (2015). How [illegible] a Big [illegible] [illegible].

[illegible] J. (2016). [illegible] Big [illegible] End [illegible] [illegible]

[illegible] B. M., [illegible] J. A. (2015). [illegible] *Journal of Sports Economics*, 16(4):353374.

[illegible] J. [illegible] P. (1985). *The Hidden Game of Baseball*. [illegible]

[illegible] D. (2008). [illegible] 2008/4/5/389840/[illegible]

[illegible] J. (2007). [illegible]

[illegible] B. (2014). [illegible]

A Model Comparison with Cross-Validation

[illegible]

	Model 1	Model 2	Model 3	Model 4	Model 5	
	# Parameters	744	855	2582	11,067	171,168
Overall	MISS	0.104	0.101	**0.101**	0.107	0.101
	MSE	0.073	0.071	**0.071**	0.072	0.072
Region 1	MISS	0.251	**0.239**	0.240	0.242	0.240
	MSE	0.164	0.158	**0.157**	0.159	0.158
Region 2	MISS	0.213	0.208	**0.206**	0.208	0.207
	MSE	0.154	0.150	**0.148**	0.150	0.149

B Catcher and Count Effects

[illegible]

[illegible]

−2 [illegible] 1 −2 [illegible]

[illegible] 3 −1 [illegible] 1 −0 [illegible]

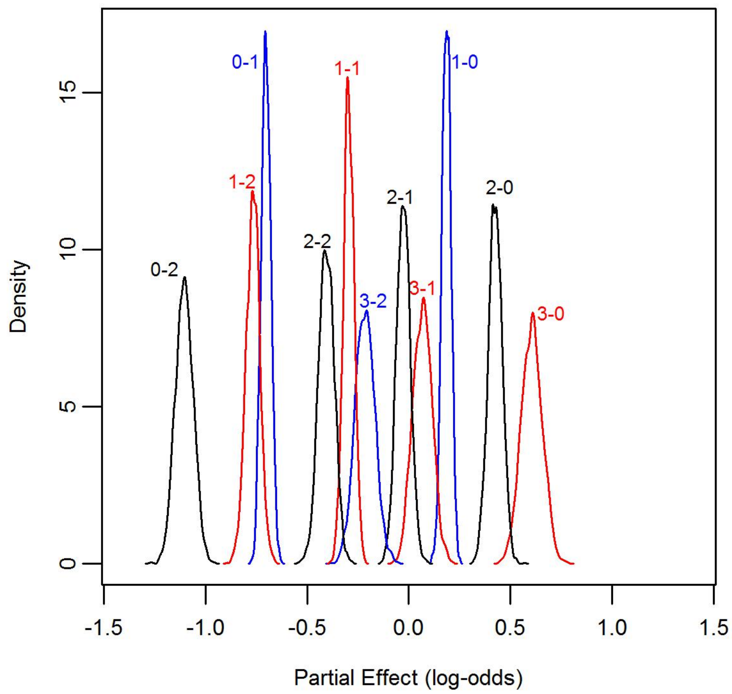
[illegible] 3 −0 [illegible]

[illegible] 0 −1 [illegible]

[illegible]

[illegible]

[illegible] 0 −1 [illegible] 3 −0 [illegible] (37,513 [illegible] 6,162).



Fig[illegible]7: P[illegible] D[illegible]
[illegible]

[illegible]

[illegible] P[illegible] B[illegible] P[illegible] M[illegible]

B[illegible] B[illegible] GAM [illegible]

[illegible] 50%. A[illegible]

[illegible] B[illegible] P[illegible] 93 [illegible] 54%, [illegible] 64% [illegible]

I[illegible], [illegible] B[illegible] P[illegible]

[illegible] 64%, [illegible] P[illegible] 10% [illegible]

[illegible] If P[illegible] 2 –

2 [illegible] 0 – 0, [illegible] 55%. I[illegible]

[illegible] P[illegible] 0 – 0 [illegible]

[illegible] 0 – 0 [illegible] 2 – 2 [illegible] P[illegible]

[illegible]

[illegible]

[illegible] B[illegible] P[illegible] GAM

[illegible] 50% [illegible]

[illegible] 54%, [illegible]

[illegible] P[illegible] 0 – 0, [illegible]

[illegible]: Diff[illegible] B[illegible] P[illegible] 54%, [illegible] C[illegible] in Fig[illegible] 7

	Hank Conger	Buster Posey	Brayan Pena	Tomas Telis
0 – 2	-0.14	-0.18	-0.26	-0.39
1 – 2	-0.06	-0.10	-0.18	-0.35
0 – 1	-0.04	-0.08	-0.17	-0.34
2 – 2	0.03	-0.01	-0.10	-0.28
1 – 1	0.06	0.02	-0.07	-0.26
3 – 2	0.07	0.0 4	-0.05	-0.25
2 – 1	0.12	0.08	-0.01	-0.20
0 – 0	0.12	0.09	0.00	-0.20
3 – 1	0.14	0.10	0.02	-0.18
1 – 0	0.16	0.13	0.04	-0.16
2 – 0	0.21	0.18	0.10	-0.10
3 – 0	0.24	0.21	0.14	-0.06

[illegible] 54% [illegible] A[illegible],

[illegible] 0 – 0 [illegible] 0 – 2 [illegible] P[illegible]

[illegible] 0 – 0 [illegible] 1 – 2 [illegible]

[illegible] H[illegible]

C[illegible] P[illegible] F[illegible]

[illegible] 15% [illegible] 0 –2 [illegible]

[illegible] 40% [illegible] 36% [illegible] 28% [illegible]